Roadways, Input Sourcing, and Patterns of Specialisation

Esteban Jaimovich*

January 2018

Abstract

We propose a model where the internal transport network facilitates the sourcing of intermediate goods from different locations. Efficient internal transportation promotes thus the growth of industries that rely on a large variety of inputs. The model shows that heterogeneities in internal transport infrastructures can become a key factor in shaping comparative advantage and specialisation. Moreover, when sufficiently pronounced, such heterogeneities may even overshadow more traditional sources of specialisation based on factor productivities. Evidence based on industry-level trade data grants support to the main prediction of the model: countries with denser road networks export relatively more in industries that exhibit wider input bases. We show that this correlation is robust to several possible confounding effects proposed by the literature, such as the impact of institutions on specialisation in complex goods. Furthermore, we show that a similar correlation also arises when the density of the local transport network is measured by the density of their internal waterways, rather than by roadway density.

Keywords: International Trade, Comparative Advantage, Internal Transportation Costs.

JEL Classifications: F11, O18, R12

*University of Surrey. Mailing address: School of Economics, Guildford, GU2 7XH, United Kingdom. Email: e.jaimovich@surrey.ac.uk

1 Introduction

The spatial distribution of economic activities means that transportation costs represent a major factor influencing countries' output, trade flows and specialisation. Apart from few exceptions, the vast majority of the past trade literature has centred their attention on the cost of shipping goods internationally.[1] However, the evidence at hand suggests that internal transport costs are far from being a secondary component that can be disregarded when confronted with transboundary costs.[2] Furthermore, the impact of internal transport costs on specialisation gets magnified by the fact that local infrastructures differ quite substantially between countries, especially when comparing economies at different stages of development.

Being able to efficiently transport commodities across space is crucial to keep total costs low. Yet, owing to specificities of their physical characteristics and of their production processes, some commodities turn out to be inherently more transport-intensive than others. This means that the efficiency of the local transportation infrastructure may unevenly affect the development of different industries. This paper studies a specific channel by which the internal transport network may shape countries' comparative advantages and specialisation. We argue that one key role of the internal transportation network is that it facilitates the sourcing of intermediate inputs from different locations. As a result, industries that require a large variety of intermediate inputs tend to make more intense use of the network.

To illustrate this idea, we introduce a simple model with two intermediate inputs and a continuum of final good producers. A denser road network allows cheaper transportation of the intermediate inputs to the location site of final good producers. A crucial feature of the model is that industries producing final goods differ in terms of the breadth of their intermediate input requirements. In particular, some industries have production functions that are very intensive in only one intermediate input, while others require a more balanced use of the two intermediate inputs. Since transportation of inputs is costly, those industries that require a relatively balanced combination of the intermediate inputs turn out to benefit relatively more (in terms of cost reduction) from a denser road network.

This simple mechanism yields a very clear prediction in terms of specialisation when it is incorporated into an international trade model. Countries that enjoy a denser local trans-

[1] For a few papers that have incorporated internal transport costs into trade models, see Allen and Arkolakis (2014), Coşar and Fajgelbaum (2016), Ramondo, Rodriguez-Clare and Saborio-Rodriguez (2016), Redding (2016), Matsuyama (2017).

[2] See, e.g., Limao and Venables (2001), Anderson and Van Wincoop (2004), Hillberry and Hummels (2008), Mesquita Moreira *et al* (2013), Agnosteva *et al* (2014), Atkin and Donaldson (2015), Donaldson (forthcoming).

port network tend to display a comparative advantage in the goods whose production process requires a relatively balanced mix of the intermediate inputs. This is because these are the industries that make heavier use of the local transport network to source their inputs. Conversely, countries with underdeveloped transport networks tend to specialise in industries with narrow input bases, as this allows them to economise on input sourcing.

After presenting the model we provide evidence consistent with its main prediction. To do so, we proceed as follows. Firstly, we index industries by their degree of input breadth using the information contained in the US input-output matrix. Secondly, we measure the density of local transport networks of countries by the length of their roadways per square kilometer. Finally, we correlate countries specialisation by industries (measured by their total exports at the industry level) with an interaction term between industries' input breadth and countries' roadways density. We find that countries with denser road networks export relatively more in industries that exhibit a wider input base.

The correlation between road density and specialisation in industries with wider input bases may obviously be driven by other mechanisms to the one suggested by our model. We show however that this correlation is robust to the inclusion of a large set of possible confounding covariates. In particular, one important channel related to ours works through institutions, as industries that rely on a wide set of inputs tend to be more dependent on contract enforcement [Levchenko (2007) and Nunn (2007)]. We show that the correlation predicted by our model is still present once we also control for the effect of institutions. In that respect, our findings complement the previous studies that have interpreted the degree of input variety as a sign of product complexity, showing that industries with wide input bases seem also to be strongly reliant on the internal transport network.

One additional concern is whether the found correlation can be interpreted at all as evidence of *causation* from road density to specialisation in transport-intensive industries. Roadways are the result of investment choices. Hence, road infrastructure may positively respond to transport needs resulting from patterns of specialisation, reversing thus the direction of causation. Interestingly, we show that an analogous correlation to that one found with road density arises when using *waterways* density as an alternative measure of the depth of the local transport network. Moreover, this correlation is especially strong and significant in the case lower-income countries, which are exactly the types of economies that tend to suffer from thinner road networks. The evidence based on waterways density does not directly address the possibility of a reverse causation bias present when using length of roadways. However, since waterways cannot be molded and expanded as flexibly as road networks, their results are certainly much less sensitive to issues of reverse causation than those obtained using roadways.

There is a growing literature studying the impact of the local transport infrastructure on international and intra-regional trade and specialisation. For example, Volpe Martincus and Blyde (2013) study the access to foreign markets and international trade across regions in Chile, Coşar and Demir (2016) does so for Turkey, and Volpe Martincus, Carballo and Cusolito (2017) for Peru. Donaldson (forthcoming) looked at reductions of price and output distortions across Indian regions after expansions of the local railroad network, and Donaldson and Hornbeck (2016) assess how the expansion of the railroad network in the US enhanced market access of US counties. Fajgelbaum and Redding (2014) and Coşar and Fajgelbaum (2016) investigate the regional location of export-oriented activities given the local infrastructure in the cases of Argentina and China, respectively. Closer to our main focus, Duranton, Morrow and Turner (2014) and Coşar and Demir (2016) have tried to capture whether there is some effect of road infrastructure on specialisation in transport-intensive activities. Duranton *et al* (2014) show that US cities with more highways tend to produce goods of higher weight per physical unit, while Coşar and Demir (2016) find a similar effect for Turkey. We focus on different channel whereby the local transport infrastructure may impact comparative advantages: the notion that the spatial distribution of activities makes industries that need to source a large variety of intermediate inputs relatively more reliant on the internal transport network. Furthermore, our paper looks a cross-country variation in transport networks. As a result, it can make contact with the traditional trade literature by addressing questions related to comparative advantage and specialisation at the country level. The above-mentioned literature instead restricts the analysis to the regional distribution of activities within single economies, implying that they focus on contexts that tend to be characterised by free (internal) labour mobility, and hold fixed the level of institutions, financial development and other country-level features that do vary when looking at different economies.

Our paper also relates to two recent strands of literature have expanded upon the traditional Ricardian/Heckscher-Ohlin trade models based on factor endowments. One set of papers have looked at institutions as a source of comparative advantage in industries producing complex goods that require a large variety of input-specific relationships [Antràs (2005), Acemoglu, Antràs and Helpman (2007), Levchenko (2007), Nunn (2007), Costinot (2009)]. The other one, has delved into the role of financial markets fostering exports in industries that are heavy users of external finance [Beck (2002), Svaleryd and Vlachos (2005), Becker, Chen and Greenberg (2012), Manova (2013)]. Our paper seeks to highlight the impact of local infrastructure when industries differ in how dependent they are on internal transportation of inputs.

The rest of the paper is organised as follows. Section 2 introduces the main features of the model in the case of a closed economy. Section 3 extends the model to a two-country setup, and derives the main predictions in terms of comparative advantage and trade flows. Section 4 contrasts the main predictions of the model with the data. Section 5 discusses some endogeneity issues and alternative interpretations of the empirical results. Section 6 concludes.

2 General Setup in a Closed Economy Model

This section presents the environment and main features of our model in the specific case of a closed economy. Starting with a closed economy proves helpful in two aspects. First, it allows an easier description of the main building blocks of the model. Second, it facilitates the task of providing the main intuition for how the density of the transport network may heterogeneously affect the cost of production in different sectors.

2.1 Intermediate and Final Goods Sector

There exists a unit continuum of final goods, indexed by the letter $j \in [0, 1]$. All final good markets are perfectly competitive. Final goods are purchased by individuals with preferences given by

$$U = \int_0^1 \ln(y_j)\, dj, \tag{1}$$

where y_j denotes the consumed amount of j. There is a mass of individuals equal to L. Each individual is endowed with one unit of labour which is supplied inelastically for a wage w.

In addition to the set of final goods, there exist two intermediate goods, indexed by $i = 0, 1$. The markets for both intermediate goods are perfectly competitive. Each intermediate good is produced with labour, according to the following linear production functions:

$$X_i = \frac{L_i}{1 + \varepsilon_i}, \qquad i = 0, 1. \tag{2}$$

In (2), X_i denotes the total amount of intermediate good i produced in the economy, L_i is the total amount of labour used in producing i, and $\varepsilon_i \geq 0$ is a technological parameter determining labour productivity in sector i.

Final goods are produced by combining the two intermediate goods within Cobb-Douglas production functions. Total output of final good $j \in [0, 1]$ is given by:

$$Y_j = \frac{1}{\alpha_j^{\alpha_j} (1 - \alpha_j)^{1-\alpha_j}}\, X_{0,j}^{1-\alpha_j}\, X_{1,j}^{\alpha_j}, \qquad \text{where } \alpha_j \in [0, 1], \tag{3}$$

and $X_{0,j}$ and $X_{1,j}$ denote the amount of intermediate good 0 and 1 used in the production of final good j, respectively.

The Cobb-Douglas production functions (3) differ across final good sectors in terms of the intensity requirements of each intermediate good. Sectors with a small (resp. large) α_j use input 0 (resp. input 1) more intensively. On the other hand, sectors whose α_j lies in the vicinity of 0.5 tend to use a relatively balanced bundle of both inputs. For the rest of the paper, we will assume that, when considering the whole set of final good producers, the values of α_j are uniformly distributed within the unit interval. Abusing a bit the notation, we can thus henceforth index final goods by their value of α_j.

Perfect competition in final good markets implies that, in equilibrium, each final good j will be sold at a price equal to its marginal cost. Using (3), we can obtain the expression for the marginal cost, which we denote by c_j. Namely,

$$c_j = p_{0,j}^{1-\alpha_j}\, p_{1,j}^{\alpha_j}, \tag{4}$$

where $p_{0,j}$ and $p_{1,j}$ are the prices at which the producer of final good j can purchase each unit of input 0 and 1, respectively.[3]

2.2 Geographic Structure of the Economy

We assume that each intermediate good is produced in a different site, which we refer to as location 0 (for input 0) and location 1 (for input 1). Labour is perfectly mobile across locations at zero cost. Intermediate goods must, however, incur in an iceberg transport cost to be moved around. In particular, when the distance between the location of j and that of i is $d_{j,i} \geq 0$, the intermediate good producer i must ship $1 + t\, d_{j,i}$ units of input i in order for the final good producer j to receive one unit of i.

There exists a road network of length r linking location 0 and location 1. We assume that the shortest distance between location 0 and 1 is given by a function $\varphi(r)$, with $\varphi'(r) < 0$. That is, we assume that longer road networks facilitate transportation across location 0 and 1 by shortening the distance between the two locations. In Appendix A, we provide a simple geographical structure of the economy as microfoundation of the function $\varphi(r)$ and the fact that is strictly decreasing in r.

[3] Although we are assuming that intermediate goods are sold in competitive markets, in principle, our model will not always lead to the same price paid by each final good producer j for each of the inputs. The reason for this is that both $p_{0,j}$ and $p_{1,j}$ will also incorporate internal transport costs, and these costs may well differ across final good producers given their location and the locations of intermediate goods.

2.3 Location Choice by Final Good Producers

The previous subsection assumed that each intermediate good is produced in a specific and exogenously given location. With regards to final goods producers, we assume that they can freely choose a location on any point along the road network linking location 0 and location 1.

Given that shipping inputs across production sites entails a transport cost, final good producers will choose their own location so as to minimise their marginal costs (c_j). Recall that, given a road network of length r, the distance between location 0 and 1 is equal to $\varphi(r)$. Let now $l_j\varphi(r)$ denote the (minimum) distance between the location chosen by producer j and location 0, where $l_j \in [0,1]$. Notice $l_j = 0$ means that j selects location 0, while $l_j = 1$ means that j chooses location 1. On the other hand, interior values of l_j –that is, $l_j \in (0,1)$– entail that j locates itself at somewhere *along* the road network linking location 0 and 1.

Given the selected $l_j \in [0,1]$, producer j must thus pay

$$p_{0,j} = [1 + l_j\varphi(r)t]\,(1+\varepsilon_0)\,w$$

for each unit of input 0 that he purchases, while he must pay

$$p_{1,j} = [1 + (1-l_j)\,\varphi(r)t]\,(1+\varepsilon_1)\,w$$

for each purchased unit of input 1.

Bearing in mind (4), producer j will thus choose his location by solving:

$$\min_{l_j\in[0,1]} : c_j(l_j) = [(1 + l_j\varphi(r)t)\,(1+\varepsilon_0)\,w]^{1-\alpha_j}\,[(1 + (1-l_j)\,\varphi(r)t)\,(1+\varepsilon_1)\,w]^{\alpha_j}\,. \tag{5}$$

The above problem yields corner solutions. Comparing thus $c_j(0)$ vis-a-vis $c_j(1)$, we obtain

$$l_j^* = \begin{cases} 0 & \text{if } \alpha_j \leq 0.5 \\ 1 & \text{if } \alpha_j \geq 0.5 \end{cases} \tag{6}$$

The result in (6) is quite intuitive: final producers choose to locate their firm in the same place where the input they use more intensively is being produced.

Finally, plugging (6) back into the expression in the right-hand side of (5) we can obtain the marginal cost of final good j:

$$c_j^* = \begin{cases} (1+\varphi(r)t)^{\alpha_j}\,(1+\varepsilon_0)^{1-\alpha_j}\,(1+\varepsilon_1)^{\alpha_j}\,w & \text{if } \alpha_j \leq 0.5 \\ (1+\varphi(r)t)^{1-\alpha_j}\,(1+\varepsilon_0)^{1-\alpha_j}\,(1+\varepsilon_1)^{\alpha_j}\,w & \text{if } \alpha_j \geq 0.5 \end{cases} \tag{7}$$

The expressions in (7) shows that the marginal cost of final good j is determined by the labour cost of producing the required inputs (via the wage w, and the parameters ε_0 and ε_1), and *also* by the transport cost involved in sourcing those inputs. Importantly, recall that final good producers will optimally choose to set up their firms in the *same* location where the input they use more intensively is being produced. As a result, the transport cost ends up being applied *only* to the input whose Cobb-Douglas weight in (3) is *smaller* than 0.5. In turn, this implies that internal transport costs tend to affect more severely the marginal cost of those final goods whose α_j lies near 0.5. In other words, internal transport costs tend to particularly hurt sectors which use a relatively balanced combination of inputs. On the other hand, this also implies that while improvements in transport infrastructure will lower the cost of production of all final goods (except for the extreme cases where either $\alpha_j = 0$ or $\alpha_j = 1$), such improvements will end up lowering the marginal cost of goods whose α_j is closer to 0.5 by relatively more. The following lemma states this result more formally.

Lemma 1 *Consider the expression for the marginal cost of good j in (7) and two generic values of the road length r_1 and r_2, such that $r_1 < r_2$. Then,*

1. $c_j^*(r_1)/c_j^*(r_2) > 1$ *for all* $\alpha_j \in (0,1)$*, while* $c_j^*(r_1)/c_j^*(r_2) = 1$ *when* $\alpha_j = 0$ *and* $\alpha_j = 1$.

2. *The ratio* $c_j^*(r_1)/c_j^*(r_2)$ *is strictly increasing in* α_j *for all* $\alpha_j \in \left[0, \frac{1}{2}\right)$ *and strictly decreasing in* α_j *for all* $\alpha_j \in \left(\frac{1}{2}, 1\right]$. *Moreover, the highest value of* $c_j^*(r_1)/c_j^*(r_2)$ *is reached when* $\alpha_j = \frac{1}{2}$.

Lemma 1 shows that larger values of r lead to lower marginal costs of production, but that the fall in the marginal cost is proportionally greater in sectors with values of α_j close to $\frac{1}{2}$. In the next section, where we extend the model to allow international trade and specialisation, this result will turn the density of the road network into a source of a comparative advantage. More precisely, countries with denser road networks will tend enjoy a comparative advantage in sectors with intermediate levels of α_j.

3 Two-Country Model

We consider now a world economy a lá Dornbusch-Fischer-Samuelson (1977) with two countries: H and F. Both countries are populated by a mass L of individuals. Each individual is endowed with one unit of labour that is supplied inelastically in the local labour market. We let w_H and w_F denote the wage in H and F, respectively. Henceforth, we set $w_F = 1$ (i.e., we set w_F as the *numeraire*), and use $\omega \equiv w_H/w_F$ to denote the relative wage. All individuals share the same preferences –given by (1)– over the unit continuum of final goods.

Each final good could in principle be produced by any of the two countries. The technologies to produce final goods are identical in both H and F, given by the Cobb-Douglas functions (3). All final goods markets are perfectly competitive. In addition, we assume that all final goods are internationally tradeable, subject to an iceberg cost $\tau > 0$ (that is, when $1 + \tau$ units of j are shipped internationally, only 1 unit of j will arrive at the destination country).

Unlike for final goods, we assume that intermediate goods are non-tradeable internationally. We also assume that the technologies to produce the intermediate goods differ between H and F. Letting $X_{i,c}$ denote the total amount of intermediate good i produced in country c, we assume that in H

$$X_{0,H} = L_{0,H} \quad \text{and} \quad X_{1,H} = \frac{L_{1,H}}{1+\varepsilon}, \tag{8}$$

while in F,

$$X_{0,F} = \frac{L_{0,F}}{1+\varepsilon} \quad \text{and} \quad X_{1,F} = L_{1,F}, \tag{9}$$

where $L_{i,c}$ is the total amount of labour used in producing input i in country c, and $\varepsilon > 0$. The intermediate goods markets are perfectly competitive both in H and in F.

Two features implied by (8) and (9), coupled with the final goods production functions (3), are worth stressing here. First, since they imply that H is relatively more productive than F in the intermediate sector 0, they tend to yield a comparative advantage by H on the final goods that rely more heavily on input 0 (that is, on those j whose α_j is small). Second, since (8) and (9) exactly mirror one another, they implicitly assume away any aggregate absolute advantage by one country over the other one stemming from the distribution of sectoral labour productivities.[4]

Analogously to the closed economy setup in Section 2.2, we assume that each input is produced in a specific location. We keep referring as location 0 to the production site of input 0, and as location 1 to that one of input 1. (In this case, there is one such location in each of the countries.) Also like in the closed economy setup, we assume that the distance between location 0 and 1 in country c depends on the length of the road network in c via the distance function $\varphi(r_c)$. We also assume that the iceberg cost t per unit of distance $d_{j,i}$ travelled by input i to reach producer j is identical in H and F.

[4] None of the main results of the model crucially depend on this last feature. In fact, the model could be easily generalised to encompass intermediate production functions $X_{i,c} = L_{i,c}/(1 + \varepsilon_{i,c})$, where, $i = 1, 2$, $c = H, F$ and $\varepsilon_{i,c} \geq 0$. We deliberately choose a symmetric distribution of labour productivities, as featured by (8) and (9), only because this allows depicting the influence of road networks on the patterns of comparative advantage across H and F more cleanly.

We denote now by r_H and r_F the length of the road netword in H and F, respectively. Henceforth, we assume:

Assumption 1 $r_F < r_H$.

In our model, Assumption 1 will convey a source of comparative advantage to H in the types of goods that depend on (internal) transport of inputs more strongly. In addition, $r_H > r_F$ also implies that H can, in general, ship inputs internally at lower cost than F. This fact will, in turn, grant a source of aggregate absolute advantage by H over F.

3.1 Pricing of Final Goods in *H* and *F*

The fact that all good markets in H and F are perfectly competitive implies again that final goods will be sold at their marginal costs. Notice that this will include both the incurred internal and international transport costs. In its general form, the price of final good $j \in [0,1]$ produced in country $c = H, F$ and sold in country $m = H, F$ will be given by

$$P_{j,c}^{m} = (1 + \tau \cdot \mathbb{I}_{m \neq c}) \left[(1 + l_{j,c}\, \varphi(r_c)t)\,(1 + \varepsilon_{0,c})\right]^{1-\alpha_j} \left[(1 + (1 - l_{j,c})\, \varphi(r_c)t)\,(1 + \varepsilon_{1,c})\right]^{\alpha_j} w_c, \quad (10)$$

where: $i)$ $\mathbb{I}_{m \neq c}$ is an indicator function that is equal to one when $m \neq c$, and zero otherwise; $ii)$ $\varepsilon_{0,H} = \varepsilon_{1,F} = 0$ and $\varepsilon_{1,H} = \varepsilon_{0,F} = \varepsilon$; $iii)$ $l_{j,c}\, \varphi(r_c)$, where $l_{j,c} \in [0,1]$, is the (minimum) distance between producer j in country c and location 0.

Final good producers will optimally seek to minimise their marginal costs. Analogously as done in Section 2.3, it can be proved that this is achieved by setting up firm j in location 0 when $\alpha_j \leq 0.5$, and setting it up in location 1 when $\alpha_j \geq 0.5$. That is, condition (6) still holds true within the two-country model, with $l_{j,c} = l_j^*$ for $c = H, F$.

By using this result, together with (10), the price of good j when produced in country H and sold in $m = H, F$, denoted by $P_{j,H}^{m}$, can be written as

$$P_{j,H}^{m} = \begin{cases} (1 + \tau \cdot \mathbb{I}_{m \neq H})\ (1 + \varphi(r_H)t)^{\alpha_j}\ (1 + \varepsilon)^{\alpha_j}\ \omega & \text{if} \quad \alpha_j \leq 0.5 \\ (1 + \tau \cdot \mathbb{I}_{m \neq H})\ (1 + \varphi(r_H)t)^{1-\alpha_j}\ (1 + \varepsilon)^{\alpha_j}\ \omega & \text{if} \quad \alpha_j \geq 0.5 \end{cases}. \quad (11)$$

Analogously, $P_{j,F}^{m}$, with $m = H, F$, can be written as

$$P_{j,F}^{m} = \begin{cases} (1 + \tau \cdot \mathbb{I}_{m \neq F})\ (1 + \varphi(r_F)t)^{\alpha_j}\ (1 + \varepsilon)^{1-\alpha_j} & \text{if} \quad \alpha_j \leq 0.5 \\ (1 + \tau \cdot \mathbb{I}_{m \neq F})\ (1 + \varphi(r_F)t)^{1-\alpha_j}\ (1 + \varepsilon)^{1-\alpha_j} & \text{if} \quad \alpha_j \geq 0.5 \end{cases}. \quad (12)$$

To ease notation, it proves convenient to define

$$\delta \equiv \frac{1 + \varphi(r_F)t}{1 + \varphi(r_H)t}. \quad (13)$$

Notice that $\delta > 1$, since $r_F < r_H$. In the context of our model, δ can be interpreted as a measure of the advantage of H over F in terms of length of road network.

3.2 Traded (and Non-Traded) Goods

In equilibrium, consumers will buy each final good j from the producer who can offer j at the lowest price. In some cases this will mean that consumers will source good j locally, while in others they will choose to import it. Naturally, given that shipping final goods internationally entails an iceberg cost $\tau > 0$, if in equilibrium country c is an exporter of good j, then it must also be the case that individuals from c must be buying good j from local producers.

By comparing (11) vis-a-vis (12), we can observe that international trade of final goods takes place when the following conditions hold true (henceforth, without any loss of generality, we assume that when confronted with identical prices, consumers always buy from local producers).

- H will export final good j to F if and only if:

$$\begin{aligned} &\omega < (1+\tau)^{-1}\,\delta^{\alpha_j}\,(1+\varepsilon)^{1-2\alpha_j} && \text{when } \alpha_j \leq 0.5, \\ &\omega < (1+\tau)^{-1}\,\delta^{1-\alpha_j}\,(1+\varepsilon)^{1-2\alpha_j} && \text{when } \alpha_j \geq 0.5 \end{aligned} \tag{14}$$

- H will import final good j from F if and only if:

$$\begin{aligned} &\omega > (1+\tau)\,\delta^{\alpha_j}\,(1+\varepsilon)^{1-2\alpha_j} && \text{when } \alpha_j \leq 0.5, \\ &\omega > (1+\tau)\,\delta^{1-\alpha_j}\,(1+\varepsilon)^{1-2\alpha_j} && \text{when } \alpha_j \geq 0.5 \end{aligned} \tag{15}$$

The presence of $\tau > 0$ in (14) and (15) implies that some final goods may end up *not* being traded internationally. In particular, if for some subset of final goods whose $0 \leq \alpha_j \leq 0.5$, the model yields $(1+\tau)^{-1} \leq \omega\delta^{-\alpha_j}(1+\varepsilon)^{2\alpha_j-1} \leq (1+\tau)$, then consumers from both H and F will end up sourcing these goods locally. Similarly, if for some subset of final goods whose $0.5 \leq \alpha_j \leq 1$, the model yields $(1+\tau)^{-1} \leq \omega\delta^{\alpha_j-1}(1+\varepsilon)^{2\alpha_j-1} \leq (1+\tau)$, these goods will also be sourced in both H and F from local producers.

3.3 Equilibrium and Patterns of Specialisation

In equilibrium, the total (world) spending on final goods produced in each country must equal the total labour income of each country. In our two-country setup, this condition can be restated as a trade balance equilibrium for either H or F. The utility function (1) implies that consumers allocate identical expenditure shares across all final goods in the optimum.[5] Hence, in our model, the equilibrium condition in the world economy boils down to:

$$\int_0^{\frac{1}{2}} \mathbb{I}\left\{\omega < (1+\tau)^{-1}\delta^{\alpha_j}(1+\varepsilon)^{1-2\alpha_j}\right\} d\alpha_j + \int_{\frac{1}{2}}^{1} \mathbb{I}\left\{\omega < (1+\tau)^{-1}\delta^{1-\alpha_j}(1+\varepsilon)^{1-2\alpha_j}\right\} d\alpha_j =$$
$$\left[\int_0^{\frac{1}{2}} \mathbb{I}\left\{\omega > (1+\tau)\delta^{\alpha_j}(1+\varepsilon)^{1-2\alpha_j}\right\} d\alpha_j + \int_{\frac{1}{2}}^{1} \mathbb{I}\left\{\omega > (1+\tau)\delta^{1-\alpha_j}(1+\varepsilon)^{1-2\alpha_j}\right\} d\alpha_j\right]\omega, \tag{16}$$

where $\mathbb{I}\{\cdot\}$ in (16) is an indicator function that is equal to 1 when the condition inside the parenthesis holds true, and 0 otherwise. The left-hand side of (16) thus amounts to the total value of H's exports, whereas its right-hand side equals the total value of H's imports.

Henceforth, we impose an additional parametric restriction to the model:

Assumption 2 $\varepsilon > \tau$.

Assumption 2 ensures that our model will always feature positive trade in equilibrium. Intuitively, $\varepsilon > \tau$ implies that the source of comparative advantages linked to heterogeneities in sectoral labour productivities –i.e., those determined by (8) and (9)– are strong enough so as not to be completely overturned by international trade costs in all final sectors.[6]

From the trade balance equilibrium condition (16) we can obtain our first result concerning the equilibrium relative wage, ω^*.

Proposition 1 *In equilibrium, the wage in H is strictly greater than in F. That is, $\omega^* > 1$. Furthermore, ω^* is strictly increasing in δ, and $\omega^* < \min\left\{(1+\tau)\delta^{\frac{1}{2}}, (1+\varepsilon)(1+\tau)^{-1}\right\}$ if $(1+\varepsilon)^2 > \delta$, whilst $\omega^* < (1+\tau)^{-1}\delta^{\frac{1}{2}}$ if $(1+\varepsilon)^2 \leq \delta$.*

The result $\omega^* > 1$ is a straightforward implication of the fact that Assumption 1 conveys an aggregate advantage by H over F. As a result, in equilibrium, ω must rise above one, in order to allow F to be able to export to H as much as H exports to F. Notice that since labour is the only non-reproducible input in our model, wages are also equal to income per head in each country. Thus, Proposition 1 is ultimately stating that H is richer than F.

[5] All the results in this section can easily be extended to a general Cobb-Douglas utility function with constant (but non-equal) expenditure shares across goods. The specific choice of (1) is just for algebraic simplicity.

[6] Assumption 2 is a *sufficient* condition (but is not a *necessary* condition) to ensure that positive trade between H and F always takes place in equilibrium. Intuitively, Assumption 1 creates another source of comparative advantage in our model, in addition to heterogeneities in sectoral labour productivities. As a result, even when $\varepsilon \leq \tau$, our model may still deliver positive trade, provided δ is sufficiently large.

For future reference it proves convenient to define four different thresholds for α_j, namely:

$$\underline{\alpha}_H \equiv \frac{\ln(1+\varepsilon)-\ln(1+\tau)-\ln(\omega^*)}{2\ln(1+\varepsilon)-\ln(\delta)} \tag{17}$$

$$\overline{\alpha}_H \equiv \frac{\ln(1+\varepsilon)-\ln(1+\tau)+\ln(\delta)-\ln(\omega^*)}{2\ln(1+\varepsilon)+\ln(\delta)} \tag{18}$$

$$\underline{\alpha}_F \equiv \frac{\ln(1+\varepsilon)+\ln(1+\tau)-\ln(\omega^*)}{2\ln(1+\varepsilon)-\ln(\delta)} \tag{19}$$

$$\overline{\alpha}_F \equiv \frac{\ln(1+\varepsilon)+\ln(1+\tau)+\ln(\delta)-\ln(\omega^*)}{2\ln(1+\varepsilon)+\ln(\delta)}. \tag{20}$$

The above thresholds are obtained from the expressions in (11) and (12) in the following way: $\underline{\alpha}_H$ solves $P^F_{j,F}(\underline{\alpha}_H) = P^F_{j,H}(\underline{\alpha}_H)$ and $\underline{\alpha}_F$ solves $P^H_{j,H}(\underline{\alpha}_F) = P^H_{j,F}(\underline{\alpha}_F)$ when $\alpha_j \leq 0.5$, whereas $\overline{\alpha}_H$ solves $P^F_{j,F}(\overline{\alpha}_H) = P^F_{j,H}(\overline{\alpha}_H)$ and $\overline{\alpha}_F$ solves $P^H_{j,H}(\overline{\alpha}_F) = P^H_{j,F}(\overline{\alpha}_F)$ when $\alpha_j \geq 0.5$. Hence, the thresholds $\underline{\alpha}_H$ and $\overline{\alpha}_H$ (resp. $\underline{\alpha}_F$ and $\overline{\alpha}_F$) pin down the final goods such that, given the value of ω^*, their market price when sold in F (resp. when sold in H) would be identical regardless of where it was originally produced. Notice that $\tau > 0$ implies $\overline{\alpha}_H < \overline{\alpha}_F$, while Assumption 2 together with the equilibrium result $\omega^* > 1$ means that $\overline{\alpha}_F < 1$. Furthermore, $\underline{\alpha}_H < \underline{\alpha}_F$ when $(1+\varepsilon)^2 > \delta$, while $\underline{\alpha}_H > \underline{\alpha}_F$ holds when $(1+\varepsilon)^2 < \delta$. In addition, the results in Proposition 1 concerning the bounds for ω^* imply that $\underline{\alpha}_H > 0$ always holds.[7]

By using the thresholds (17)-(20), we can fully split the space of final goods according to their price in the destination country, given the country of origin of the good.

Lemma 2 *From (11) and (12), and the equilibrium relative wage ω^*, by using (17)-(20), we can derive the following set of conditions for $P^F_{j,F}$ relative to $P^F_{j,H}$ and for $P^H_{j,H}$ relative to $P^H_{j,F}$:*

1. *Suppose $\delta < (1+\varepsilon)^2$, then:*

   - $P^F_{j,H} < P^F_{j,F}$ *for* $0 \leq \alpha_j < \alpha^*_H$, *while* $P^F_{j,H} > P^F_{j,F}$ *for* $\alpha^*_H < \alpha_j \leq 1$, *where* $\alpha^*_H = \underline{\alpha}_H$ *if* $\omega^* \geq (1+\tau)^{-1}\delta^{\frac{1}{2}}$ *holds in equilibrium, while* $\alpha^*_H = \overline{\alpha}_H$ *if instead* $\omega^* < (1+\tau)^{-1}\delta^{\frac{1}{2}}$.
   - $P^H_{j,H} < P^H_{j,F}$ *for* $0 \leq \alpha_j < \overline{\alpha}_F$, *while* $P^H_{j,H} > P^H_{j,F}$ *for* $\overline{\alpha}_F < \alpha_j \leq 1$.

[7] The comparisons of $\underline{\alpha}_H$ vis-a-vis $\overline{\alpha}_H$ and $\underline{\alpha}_F$ vis-a-vis $\overline{\alpha}_F$ are somewhat more convoluted, as they involve several possible combinations of parametric configurations and feasible solutions for ω^* given those configurations. For example, whenever $(1+\varepsilon)^2 < \delta$ holds true, for any feasible values of ω^*, we have $0 < \underline{\alpha}_H < 0.5 < \overline{\alpha}_H < 1$ and $\underline{\alpha}_F < 0.5 < \overline{\alpha}_F < 1$. Instead, when $(1+\varepsilon)^2 > \delta$, we have $\underline{\alpha}_H < \overline{\alpha}_H$ iff $\omega^* > (1+\tau)^{-1}\delta^{\frac{1}{2}}$, and $\underline{\alpha}_F < \overline{\alpha}_F$ iff $\omega^* > [(1+\tau)\delta]^{\frac{1}{2}}$ holds in that range; both conditions fail to hold true for τ sufficiently close to zero.

2. Suppose $\delta > (1+\varepsilon)^2$, then:

- $P^F_{j,H} < P^F_{j,F}$ *for* $\underline{\alpha}_H < \alpha_j < \overline{\alpha}_H$*, while* $P^F_{j,H} > P^F_{j,F}$ *for* $0 \leq \alpha_j < \underline{\alpha}_H$ *and for* $\overline{\alpha}_H < \alpha_j \leq 1$.
- $P^H_{j,H} < P^H_{j,F}$ *for* $\max\{0, \underline{\alpha}_F\} < \alpha_j < \overline{\alpha}_F$*, while* $P^H_{j,H} > P^H_{j,F}$ *for* $\overline{\alpha}_F < \alpha_j \leq 1$ *and* $0 \leq \alpha_j < \max\{0, \underline{\alpha}_F\}$*, where* $\underline{\alpha}_F > 0$ *if and only if* $\omega^* \geq (1+\tau)(1+\varepsilon)$ *holds true.*

In equilibrium, consumers in both H and F will always buy good j from the producer who can sell it in each market at the lower price. Hence, relying on Lemma 2, we can next derive the equilibrium patterns of trade and specialisation in the two-country world economy.

Proposition 2 *The patterns of specialisation and trade differ qualitative depending on whether* $\delta > (1+\varepsilon)^2$ *or* $\delta < (1+\varepsilon)^2$.

i) Ricardian-based specialisation: *When* $\delta < (1+\varepsilon)^2$*, trade patterns and specialisation are governed by heterogeneities in labour productivities. Country* H *becomes an exporter of final goods whose* $\alpha_j \in [0, \alpha^*_H)$*, where* $\alpha^*_H = \underline{\alpha}_H$ *(resp.* $\alpha^*_H = \overline{\alpha}_H$*) if* $\omega^* \geq (1+\tau)^{-1}\delta^{\frac{1}{2}}$ *(resp.* $\omega^* < (1+\tau)^{-1}\delta^{\frac{1}{2}}$*) holds true. Country* F *becomes an exporter of the final goods whose* $\alpha_j \in (\overline{\alpha}_F, 1]$*. Final goods whose* $\alpha_j \in [\alpha^*_H, \overline{\alpha}_F]$ *are sourced locally by both* H *and* F.

ii) Transport cost-based specialisation: *When* $\delta > (1+\varepsilon)^2$*, trade patterns and specialisation are governed by road network length differences between* H *and* F*. Country* H *becomes an exporter of final goods whose* $\alpha_j \in (\underline{\alpha}_H, \overline{\alpha}_H)$*. Country* F *becomes an exporter of final good whose* $\alpha_j \in [0, \underline{\alpha}_F) \cup (\overline{\alpha}_F, 1]$ *if* $\omega^* \leq (1+\varepsilon)(1+\tau)$ *holds true, while it becomes an exporter of final goods whose* $\alpha_j \in (\overline{\alpha}_F, 1]$ *if instead* $\omega^* > (1+\varepsilon)(1+\tau)$ *holds true. When* $\omega^* \leq (1+\varepsilon)(1+\tau)$ *final goods whose* $\alpha_j \in [\underline{\alpha}_F, \underline{\alpha}_H] \cup [\overline{\alpha}_H, \overline{\alpha}_F]$ *are sourced locally by both* H *and* F*, while when* $\omega^* > (1+\varepsilon)(1+\tau)$ *this happens for those goods whose* $\alpha_j \in [0, \underline{\alpha}_H] \cup [\overline{\alpha}_H, \overline{\alpha}_F]$.

The patterns of trade and specialisation described by Proposition 2 are graphically depicted in Figure 1.[8] The upper panel plots *case i)* of Proposition 2 –i.e., $\delta < (1+\varepsilon)^2$–, while and the lower panel shows *case ii)* –i.e., $\delta > (1+\varepsilon)^2$.[9] The vertical axis of Figure 1 orders final goods according to their specific $\alpha_j \in [0,1]$; the horizontal one measures the relative wage ω.

[8] The solid line in Figure 1 is obtained by plotting $\underline{\alpha}_H$ and $\overline{\alpha}_H$, as given by (17) and (18) but replacing the specific equilibrium wage ω^* by a generic $\omega \geq 0$. Similarly, the dashed line is obtained by plotting $\underline{\alpha}_F$ and $\overline{\alpha}_F$, as given by (19) and (20) for a generic $\omega \geq 0$. Note that only the parts of the solid and dashed lines below $\alpha_j = 0.5$ follow the expressions in (17) and (19), while only the parts above $\alpha_j = 0.5$ follow (18) and (20).

[9] For brevity, the upper panel of Figure 1 shows the sub-case where $\omega^* > (1+\tau)^{-1}\delta^{\frac{1}{2}}$ –implying that H exports goods with $\alpha_j \in [0, \underline{\alpha}_H)$–, while its lower panel shows the sub-case where $\omega^* \leq (1+\varepsilon)(1+\tau)$ –entailing

Consider first the upper panel of Figure 1. Given a certain level of ω, all the goods that lie below the solid line would be exported by H, while all the goods lying above the dashed line would be exported by F. The gap in between the two curves represents the set of goods that would *not* be traded internationally. As it can be observed, the set of goods exported by H gets smaller as ω increases. Conversely, the set of goods exported by F expands with ω. At the extremes, when $\omega \leq 1/(1+\tau)(1+\varepsilon)$ all final goods would be produced in H, whereas for $\omega \geq (1+\tau)(1+\varepsilon)$ they would all be produced in F (naturally, such extreme values of ω could not possibly hold in equilibrium).

At the equilibrium wage, ω^*, final goods with $\alpha_j > \overline{\alpha}_F$ are exported by F, and those with $\alpha_j < \underline{\alpha}_H$ are exported by H.[10] Hence, H becomes an exporter of the final goods that use input 0 more intensively (i.e., low-α_j goods), while F an exporter of those which use input 1 more intensively (i.e., high-α_j goods). Intuitively, the condition $\delta < (1+\varepsilon)^2$ means that differences in road network lengths between H and F are *small* relatively to their heterogeneities in sectoral labour productivities. As a result, the labour productivity differentials in the intermediate sectors –dictated by (8) and (9)– become the leading source of comparative advantage, regulating trade flows in the model.

Consider now the lower panel of Figure 1. For values of $\alpha_j > 0.5$, this graph exhibits the same qualitative features as the one in the upper panel. In fact, the interpretation of the curves within the range $\alpha_j > 0.5$ is analogous in both graphs: given a ω, the final goods that lie below the solid line would be exported by H and those lying above the dashed line would be exported by F. The main visual differences between the graphs arise when $\alpha_j < 0.5$. Within this range, the final goods located *above* the solid line would be exported by H, whereas those located *below* the dashed line would be exported by F.[11] In turn, this case leads to a pattern of specialisation that differs quite drastically from that one in the upper panel of Figure 1. When, $\delta > (1+\varepsilon)^2$, we can observe that F ends up exporting the final goods located at the two (opposite) ends of the unit set –namely, $\alpha_j \in [0, \underline{\alpha}_F)$ and $\alpha_j \in (\overline{\alpha}_F, 1]$–, while H becomes an exporter of those in the intermediate range of α_j –namely, $\alpha_j \in (\underline{\alpha}_H, \overline{\alpha}_F)$–.

that F exports goods with $\alpha_j \in [0, \underline{\alpha}_F)$ and with $\alpha_j \in (\overline{\alpha}_F, 1]$–. In Appendix A, Figure 1 (bis) plots the other two sub-cases encompassed by Proposition 2.

[10] Notice that given the utility function (1), it must be that in equilibrium $(1-\overline{\alpha}_F) \times \omega^* = \underline{\alpha}_H$, where $(1-\overline{\alpha}_F) \times \omega^*$ equals total imports by H and $\underline{\alpha}_H$ equals total exports by H.

[11] Analogously to the upper panel of Figure 1, as ω increases, the set of goods exported by H shrinks and that one exported by F expands. The gap in between the curves represents the set of goods that are not traded internationally. Finally, for $\omega < 1/(1+\tau)(1+\varepsilon)$ all final goods would end up being produced by H, while for $\omega > (1+\tau)\delta^{0.5}$ they would all end up being produced by F.

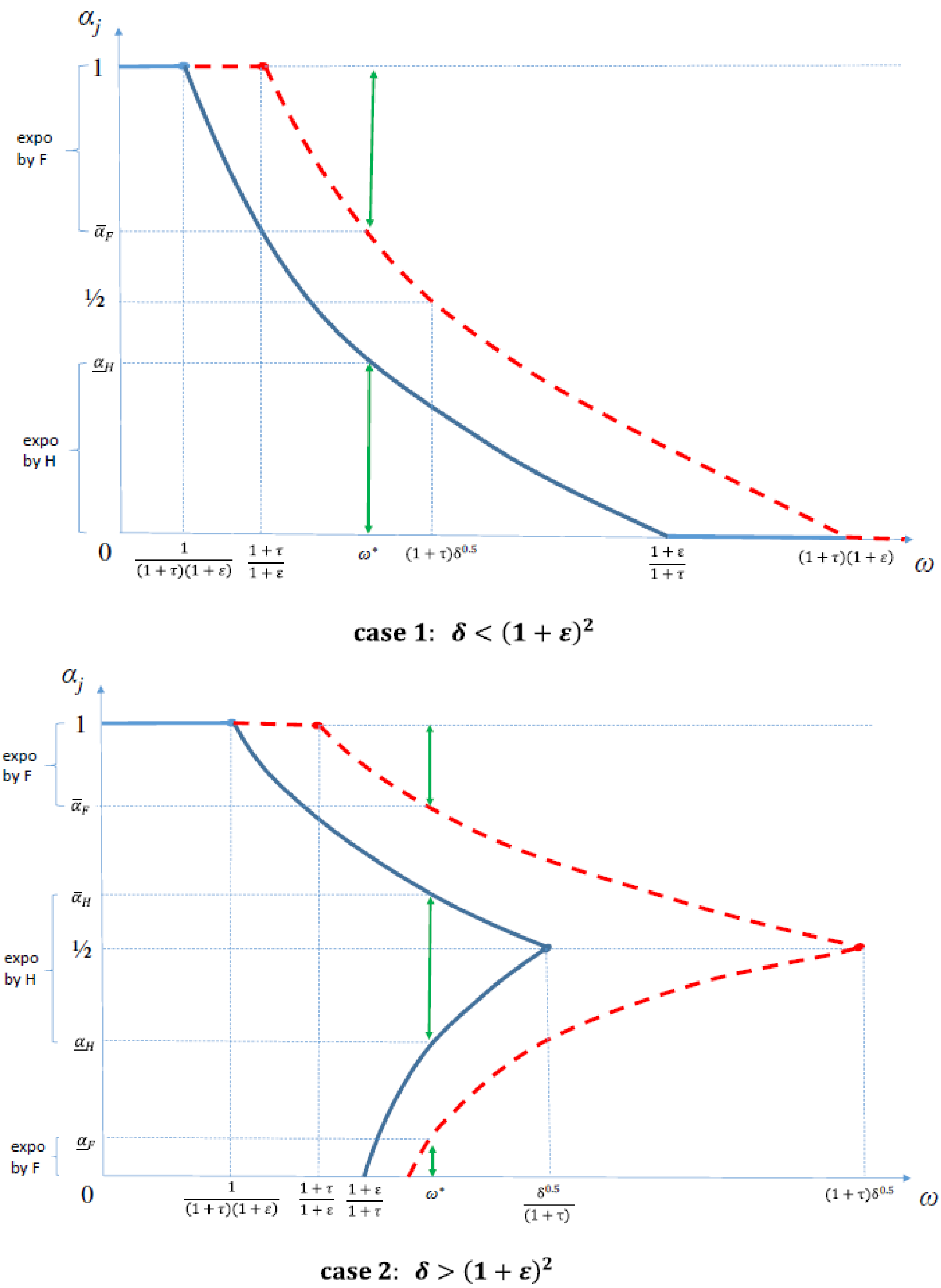


Figure 1: Patterns of Trade and Specialisation

The pattern of specialisation depicted in the lower panel of Figure 1 represents the most important insight of the model. The intuition for the result rests on the fact that when $\delta > (1+\varepsilon)^2$ the gap in the road network length is *large* relative to the heterogeneities in sectoral labour productivities, and thus becomes the leading determinant of comparative advantages. Final goods with intermediate values of α_j require the use of both inputs in similar intensity. This means that a large share of their inputs will *necessarily* have to be transported along the road network of the economy. Instead, firms producing final goods with high and low values of α_j can source a relatively large share of their inputs from the *same* location where they are located, thus without the need to rely for it on the internal transport network so heavily. In other words, sectors in the intermediate range of α_j require the use of the internal transport network more strongly than those whose α_j lies on the upper and lower spectrum of the unit interval. Accordingly, when H has a much larger road network than F, the former ends up specialising in the final goods with intermediate values of α_j, and the latter in those with more extreme values of α_j.

4 Empirical Predictions: From the Theory to the Data

In this section, we first describe how we attempt to bring to the data the main variables of interest present in the model. Next, we explain how we approach the data on trade flows to seek for evidence consistent with the main novel predictions of the model. Table A.1 in Appendix B provides some summary statistics for the main variables described below.

4.1 Main Variables of Interest

Input Narrowness

The first key issue is coming up with a measure of the breadth of set of intermediate inputs used by each industry. The model is quite stylised to allow a direct match between its technological environment and real world data on inputs and outputs by sectors. In particular, in the real world production functions tend to use more than only two intermediate goods. Furthermore, the distinction between intermediate and final goods is not so clear-cut as assumed by the model, as many goods satisfy both roles. Despite these shortcomings, we can nonetheless use the model as a guide to construct measures of narrowness of the intermediate inputs base for different industries.

In our model, sector j allocates a fraction α_j of their total spending in intermediate goods on input 0 and the remainder $(1-\alpha_j)$ on input 1. This means that sectors with very low or very high values of α_j source most of their inputs from only one intermediate sector, and thus exhibit a *narrow* intermediate input base. Conversely, sectors with values of α_j around one half rely quite importantly on both inputs, and thus display a *wide* intermediate input base.

We formally define the *narrowness* of the input base of sector j by the Gini coefficient of their expenditure shares across both inputs ($Gini_j$). The greater the value of $Gini_j$ is, the narrower input base of sector j.[12] By using the fact that expenditure shares on input 0 and 1 are given, respectively, by α_j and $1-\alpha_j$, we can observe that:

$$Gini_j = \begin{cases} \frac{1}{2}-\alpha_j & \text{if } 0 \leq \alpha_j \leq \frac{1}{2} \\ \alpha_j - \frac{1}{2} & \text{if } \frac{1}{2} \leq \alpha_j \leq 1. \end{cases} \tag{21}$$

Hence, $Gini_j = 0$ when $\alpha_j = 0.5$, while it grows symmetrically as α_j moves away from its central value of 0.5 towards either extremes on 0 and 1.

To construct a measure input narrowness analogous to that one in (21), but based on the available real world data, we resort to the input-output (IO) matrix of the US in 2007 from the Bureau of Economic Analysis (BEA).[13] The IO matrix comprises 389 sectors/industries. Although the IO matrix exhibits the same number of sectors producing intermediate goods as those producing final output, we restrict the set of final goods to those also present in the international trade data (see description below). Thus, we index by $k = 1, 2, ..., K$ each of the sectors present in the IO matrix and also in the trade data, and by $n = 1, 2, ..., N$ each of the sectors selling intermediate inputs.

We let $X_{k,n} \geq 0$ denote the total value of intermediate good n purchased by sector k. Defining $S_{k,n} \equiv X_{k,n} / \sum_{n=1}^{N} X_{k,n} \geq 0$ as the share of n over the total value of intermediates purchased by k, we can compute the Gini coefficients analogously to those in (21). Namely,

$$Gini_k = \frac{2 \times \sum_{n=1}^{N} n \times S_{k,n}}{N \times \sum_{n=1}^{N} S_{k,n}} - \frac{N+1}{N}, \tag{22}$$

where the argument $\sum_{n=1}^{N} n \times S_{k,l}$ in the numerator of $Gini_k$ is ordering intermediates in non-decreasing order (i.e., $S_{k,n} \leq S_{k,n+1}$).

[12] Imbs and Wacziarg (2003) have previously used the Gini coefficient to measure the degree of concentration of labour and value added across different sectors in the economy. In this paper, we apply a similar methodology, but we use it to measure the degree of narrowness/concentration of the intermediate input base of different sectors in the economy. There are other measures that could alternatively be used to capture the same concept; e.g., coefficient of variation, log-variance, Herfindahl index. We use those alternative measures in our empirical analysis in Section 4.3 as robustness check of the results when using the Gini.

[13] This data is publicly available from https://www.bea.gov/industry/io_annual.htm.

In the empirical analysis in Section 4.3, we use $Gini_k$ to measure the degree of narrowness of the input base of sector k. Large values of $Gini_k$ are the result of sector k sourcing most of their intermediate inputs from relatively few sectors. Conversely, small values of $Gini_k$ tend to occur when the distribution of $S_{k,n}$ is quite evenly spread across a large number of intermediates.[14] Notice finally the link between $Gini_k$ in (22) and $Gini_j$ in (21): the former boils down to the latter when $N = 2$, and $S_{k,n} = \alpha_{k,n}$ with $\alpha_{k,1} + \alpha_{k,2} = 1$.

Export Specialisation

In order to measure the degree of export specialisation by sectors we use the data on trade flows from COMTRADE compiled by Gaulier and Zignago (2010). We use only trade flows in year 2014. The data are categorised following the Harmonized System (HS) 6-digit classification, with 5,192 products. We map the trade flows data based on the HS 6-digit classification to the BEA industry codes using the concordance table between the 2002 IO matrix commodity codes and the HS 10-digit classification from the BEA website (after grouping the HS 10-digit codes into HS 6-digit products). In the cases in which an HS-6 product maps into more than one BEA code, we assign their trade flows proportionally to each of the BEA sectors in which it maps.[15] Lastly, the IO industry codes of the 2002 classification are matched to those of the 2007 classification, which are the ones actually used in the computation of the Gini coefficients.[16]

Road Network

The last key variable in our model is the length of the road network of country c (r_c). We take the road network length by countries from the data on roadways from the CIA World

[14] In the extreme (unequal) case in which $S_{k,n'} = 1$ for some n' and $S_{k,n} = 0$ for all $n \neq n'$, (22) yields $Gini_k = 2 - [(N+1)/N]$, which approaches 1 as $N \to \infty$. On the other hand, in the case when $S_{k,n} = S_k > 0$ for all $n = 1, ..., N$, we would have $Gini_k = 0$.

[15] There are 526 HS-6 products that map into two BEA Input-Output industry codes, 96 products that map into three IO codes, 33 products that map into four IO codes, and 11 products that map into five or more IO codes. (We excluded the 11 products that map into five or more IO codes.) None of the regression results in Section 4.3 are significantly altered when all the HS-6 products that map into more than one BEA Input-Output industry code are dropped from the sample.

[16] Unfortunately, we are not aware of any correspondence table between BEA 2007 codes and the HS codes, hence we indirectly link them via the BEA 2002 codes. In the end, after mapping the HS 6-digit products into the BEA 2002 codes, and mapping the BEA 2007 codes to the 2002 codes, we are left with data on trade flows and input narrowness for 294 industries as coded by the BAE 2002 classification. Of the original 389 codes, only 307 are matched to HS 10-digit codes. We have export data for 303 industries among those 307. Nine other industries are lost when matching the BAE 2007 codes to those in BAE 2002.

Factbook. Roadways are defined as ‘total length of the road network, including paved and unpaved portions’. The World Factbook contains data on roadways for 223 countries. The year of the data point for each country varies, ranging from year 1999 to 2016, with the median year of the sample being 2010. The sample we use in Section 4.3 contains 136 countries, whose measured length of roadways years range from 2000 to 2016. When defining our empirical counterpart of the variable r_c, we divide the length of the road network by the total area of the country: $r_c \equiv roadways_c/area_c$. In some of the robustness checks, we use also two additional measures of transport density: waterways density (defined as $waterways_c/area_c$) and railway density (defined as $railways_c/area_c$). The data on length of waterways and railways are also taken from the CIA World Factbook.

4.2 Road Density and Patterns of Specialisation: Testing the predictions of the model

The two-country model presented in Section 3 predicts that when heterogeneities in road networks are sufficiently large, the patterns of specialisation and trade flows follow those depicted by the lower panel of Figure 1. More formally, when the condition $\delta > (1+\varepsilon)^2$ holds true, the country with the longer road network (i.e., country H) will export goods with intermediate values of α_j, while the country with the shorter road network (i.e., country F) will export goods with values of α_j located on the extremes of the unit continuum. Conceptually, this prediction can be interpreted as stating that countries with longer road networks will tend to exhibit a comparative advantage in the types of goods that require a wider (or more diverse) set of intermediate inputs.

From an empirical viewpoint, if road network length differences across countries shaped somehow their patterns of specialisation as our model predicts, we should then observe the following: economies with a greater r_c will tend to export relatively more of the goods produced in industries with a smaller value of $Gini_k$ vis-a-vis economies with smaller r_c. We test this prediction using the following regression:

$$\ln(Expo_{k,c}) = \beta \cdot (r_c \times Gini_k) + \boldsymbol{\chi} \cdot \boldsymbol{\Delta}_{k,c} + \varsigma_c + \kappa_k + \upsilon_{c,k}, \tag{23}$$

In the regression equation (23) the dependent variable is given by the natural logarithm of the total value of exports in industry k by country c to all other countries in the world in year 2014. The term $(r_c \times Gini_k)$ interacts the measure of input narrowness defined in (22) with the measure of road density (i.e., length of roadways per square kilometer). $\boldsymbol{\Delta}_{k,c}$ denotes a vector of additional covariates that may possibly influence specialisation across countries in industries differing in terms of the degree of input narrowness. ς_c and κ_k denote country fixed effects and industry fixed effects, respectively, and $\upsilon_{c,k}$ represents an error term.

The main coefficient of interest in (23) is β. If, as the model predicts, countries with a denser road network (i.e., countries with a greater r_c) indeed tend to exhibit a comparative advantage in goods from industries that require a wider set intermediate inputs (i.e., industries with a smaller $Gini_k$), then the data should deliver a negative estimate of β.

4.3 Empirical Results

Table I displays the first set of estimation results corresponding to (23). Column (1) includes only our main variable of interest (i.e., the interaction term between r_c and $Gini_k$), together with the exporter and industry dummies. The correlation is negative and highly significant, suggesting that countries with denser road networks tend to export relatively more of the final goods whose production process requires a wider intermediate input base (i.e., those exhibiting a lower $Gini_k$). Columns (2)-(4) show the results of this simple correlation when input narrowness is measured by three alternative measures: the Herfindahl index, the coefficient of variation, and the log-variance of industry k's intermediates expenditure shares ($S_{k,n}$). The estimate of β is negative and highly significant under all these alternative measures as well. Finally, in column (5) we show the result of a regression that includes the length of the road network as a regressor, together with the interaction term $r_c \times Gini_k$. (We have to drop in this case the country fixed effects ς_c from the regression.) Since $Gini_k$ is always smaller than unity, column (5) shows that countries with longer road networks tend to export more in general, but the increase in exports is more pronounced in industries with smaller values of $Gini_k$.

In Table II we sequentially start incorporating some additional interaction terms that may also influence the patterns of specialisation across industries with different levels of input narrowness. Column (1) adds an interaction term between $Gini_k$ and an index of *Rule of Law*, taken from World Governance Indicators. The rationale behind including this term lies on the argument in Levchenko (2007) and Nunn (2007), who show that countries with better contract enforcement institutions display a comparative advantage in industries that are heavily dependent on relationship-specific investments. Within our specific context, industries that need to source a wider set of intermediate inputs may benefit relatively more from a sound legal environment, as they need to establish relationships with a greater number of input providers. Given that countries with better institutions tend to be also richer and invest more in basic infrastructure, omitting this term could lead to an overestimation (in absolute value) of the correlation coefficient of interest in (23). The regression in column (1) of Table II yields indeed

TABLE I

Export Specialisation across Industries with Different Levels of Input Narrowness

	(1)	(2)	(3)	(4)	(5)
Road Density x Input Narrowness	-7.542***	-1.455***	-0.064***	-0.225***	-8.272***
	(0.507)	(0.179)	(0.006)	(0.019)	(1.140)
Road Density					9.929***
					(1.076)
Observations	36,069	36,069	36,069	36,069	36,069
R-squared	0.755	0.754	0.754	0.754	0.299
Country FE	Yes	Yes	Yes	Yes	No
Sector FE	Yes	Yes	Yes	Yes	Yes
Number of Countries	136	136	136	136	136
Narrowness Measure	Gini	Herf	Coef Var	Log Var	Gini

Robust standard errors in parentheses. The dependent variable is the logarithm of total exports in industry k by country c in year 2014.
*** p<0.01, ** p<0.05, * p<0.1

a negative and significant coefficient associated with the interaction term between rule of law and $Gini_k$, consistent with the previous literature on institutions and specialisation. In addition to that, the absolute value of $\widehat{\beta}$ falls relative to the regression in column (1) of Table I. However, $\widehat{\beta}$ still remains negative and significant.

Another possible source of omitted variable bias is related to the effect of financial markets. There is a large body of literature that sustains that financial markets are instrumental to opening new sectors and increasing the variety of industries in the economy (e.g., Greenwood and Jovanovic, 1990; Saint-Paul, 1992; Acemoglu and Zilibotti, 1997 and 1999). We could then expect that countries with more developed financial markets would also be better able to specialise in industries that require a wider input base. To deal with this concern, in column (2) we interact the Gini coefficients with an indicator of financial development: the ratio of private credit to GDP. (This indicator is taken from the World Bank Indicators database, and averaged during years 2005-2014.) As we can readily observe, the effect of financial development interacted with $Gini_k$ is significant and it carries a sign consistent with the past literature on growth and diversification. Yet, the estimate of β still remains negative and significant.

Column (3) adds interaction terms between $Gini_k$ and GDP per capita, and between $Gini_k$ and aggregate GDP. These two terms would control for the possibility that larger or richer economies may be better able to produce goods with lower $Gini_k$, if for some reason those economies tend to be exhibit a more diversified productive structure (for example, this could happen if there are fixed costs to open some sectors). As we can see, the results concerning β remain essentially unaltered.

TABLE II

Export Specialisation across Industries with Different Levels of Input Narrowness: Additional Covariates

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Road Density x Input Narrowness	-3.881***	-3.989***	-3.933***	-2.585***	-2.472***	-10.514***	-8.639***
	(0.590)	(0.582)	(0.579)	(0.574)	(0.580)	(1.190)	(1.208)
Rule of Law x Input Narrowness	-4.625***	-1.439**	-1.544*	-1.871**	-1.810*	-1.098	-1.155
	(0.505)	(0.729)	(0.887)	(0.947)	(0.952)	(0.921)	(0.973)
Fin Dev (priv cred) x Input Narrowness		-0.095***	-0.084***	-0.047***	-0.045***	-0.068***	-0.039**
		(0.015)	(0.015)	(0.016)	(0.016)	(0.015)	(0.016)
GDP per capita x Input Narrowness			0.004	0.022	0.048	-0.004	0.016
			(0.039)	(0.039)	(0.041)	(0.039)	(0.039)
GDP x Input Narrowness			-0.906***	-0.703***	-0.711***	-0.958***	-0.756***
			(0.144)	(0.148)	(0.150)	(0.144)	(0.148)
Capital Intensity x $(K/L)_c$				0.221***			0.223***
				(0.008)			(0.080)
Skill Intensity x H_c				0.007***			0.007***
				(0.001)			(0.001)
(Pop) Density x Input Narrowness						-2.255***	-1.384***
						(0.441)	(0.471)
Road Density x (Pop) Dens x Input Nwness						2.625***	2.209***
						(0.343)	(0.346)
Observations	36,069	35,123	35,123	29,062	29,062	35,123	29,062
R-squared	0.755	0.756	0.756	0.797	0.796	0.757	0.797
Number of Countries	136	132	132	122	122	132	122
Number of Industries	294	294	294	259	259	294	259

Robust standard errors in parentheses. All regressions include country and industry fixed effects. The dependent variable is the log of exports in industry k by country c in year 2014. Rule of law is taken from the World Governance Indicators (WB) for year 2014. Private credit over GDP is taken from the World Bank Indicators, averaged for years 2005-2014. GDP per capita, GDP, stock of physical capital, and the human capital index are taken from the Penn Tables, all for year 2014. Measures of physical capital and skill intensity by industry are taken from the NBER-CES Manufacturing Industry Database, and corresponds to year 2011. *** p<0.01, ** p<0.05, * p<0.1

In column (4) we introduce two additional regressors to control for specialisation driven by factor endowments: *i*) an interaction term between capital intensity of industry k and the stock of physical capital per worker in country c, $(K/L)_c$; *ii*) an interaction term between the skill intensity of industry k and the stock of human capital in country c, H_c. The measures of capital and skill intensity at the industry level are constructed from the NBER-CES Manufacturing Industry database, for year 2011.[17] Some industries are lost from the sample when we introduce the industry factor intensity measures, since the NBER dataset contains information only for manufacturing industries. For comparability, in column (5) we display the results of the regression in column (3), but using the restricted sample. The coefficients associated to

[17] Capital intensity is computed as the total stock of physical capital per worker by industry. Skill intensity is measured by the average wage by industry. (See Becker, Gray and Marvakov (2013) for details on the NBER-CES Manufacturing Industry database.) Both the measure of physical capital per worker and the index of human capital are drawn from the Penn Tables database. (The human capital index is based on the average years of schooling from Barro & Lee (2013) and an assumed rate of return of education based on Mincer estimates.)

the factor intensities carry the expected sign, while the estimates of β remain negative and significant. Furthermore, the estimated coefficients are of similar magnitude both in columns.

Finally, the last two columns of Table II address the possibility of a differential effect of the road network on the pattern of specialisation depending on the population density of the economy. One could expect that more densely populated countries may display also a greater concentration of activities in fewer locations. Hence, all else equal, more densely populated countries may need to resort less strongly on a vast road network than sparsely populated countries do. Columns (7) and (8) assess this possibility by introducing an interaction term between population density and $Gini_k$, and a *triple* interaction term which also includes r_c. If road network length is especially important for specialisation in economies that are *less* densely populated, then the triple interaction term should carry a positive estimate. As can be observed, this is indeed the case. Moreover, the estimate of β after introducing the triple interaction term is still negative and highly significant.

Table III displays some of the regressions previously presented in Table II, but now splitting the sample of countries in two subsamples, according to whether their income is above or below the median. The odd-numbered columns show the results for the subsample of 'high-income countries', while the even-numbered columns do that for the 'low-income countries'. The results show that the effect of road density on pattern of specialisation holds true both for richer and poorer countries. In addition to that, the effect seems to be consistently greater in magnitude for the subsample of economies whose income is below the median.[18]

As additional robustness checks, Tables A.2 and A.3 in Appendix B change the measure of transport density used in the previous regressions. In Table A.2, we show the results of a set of regressions substituting r_c in (23) by railway density, computed as the total railway network length of country c by square kilometer. In Table A.3, we expand our measure transport network length to include (in addition to roadways) also the total length of internal railways and waterways. All the results in Table A.2 and A.3 follow a similar pattern as those previously shown in Table I and II.

[18] This difference in magnitude could suggest the presence of some sort of decreasing marginal effect of road density, since richer economies tend to exhibit denser road networks than poorer ones (see Figure 2 later on).

TABLE III

High-Income and Low-Income Subsamples

	(1)	(2)	(3)	(4)	(5)	(6)
Road Density x Input Narrowness	-3.656***	-7.287***	-2.458***	-4.725***	-2.398***	-4.682**
	(0.561)	(2.014)	(0.563)	(1.936)	(0.566)	(1.937)
Rule of Law x Input Narrowness	-0.986	-2.833*	-1.454	-1.954	-1.270	-1.857
	(0.958)	(1.871)	(1.019)	(2.071)	(1.018)	(2.076)
Fin Dev x Input Narrowness	-0.075***	-0.141***	-0.055***	-0.032	-0.056***	-0.035
	(0.015)	(0.051)	(0.016)	(0.055)	(0.016)	(0.055)
GDP per capita x Input Narrowness	0.006	0.348	0.037	0.587**	0.046	0.662**
	(0.043)	(0.269)	(0.044)	(0.285)	(0.045)	(0.283)
GDP x Input Narrowness	-0.893***	0.026	-0.738***	-0.007	-0.739***	0.021
	(0.139)	(0.648)	(0.142)	(0.595)	(0.145)	(0.597)
Capital Intensity x $(K/L)_c$			0.095	0.991*		
			(0.112)	(0.562)		
Skill Intensity x H_c			0.013***	-0.017***		
			(0.002)	(0.002)		
Observations	18,831	16,292	15,419	13,643	15,419	13,643
R-squared	0.765	0.623	0.805	0.653	0.805	0.651
Countries Sample (high/low income)	High	Low	High	Low	High	Low
Number of Countries	61	61	61	61	61	61
Number of Industries	294	294	259	259	259	259

Robust standard errors reported in parentheses. All regressions include country and industry fixed effects. The dependent variable is log($Expo_{c,k}$) in the year 2014. The high-income sample comprises countries with GDP per capita above the sample median, and the low-income sample countries with GDP per capita below it. The median income of the sample lies between that of Ecuador ($10,968 PPP) and Peru ($10,993 PPP) in 2014. *** p<0.01, ** p<0.05, * p<0.1

5 Beyond the Model: Endogeneity and Alternative Interpretations

The previous section presented a robust correlation between road density in country c and its degree of specialisation in industries that rely on a wide set of inputs. While those results are certainly consistent with the main predictions of the model, they cannot be taken as hard evidence of its core mechanism. In particular, two separate issues deserve some further discussion and analysis. First, the correlation found in the previous regressions could as well be the result of road infrastructure responding to transport needs stemming from industry specialisation (i.e., reverse causation). Second, our interpretation of a lower value of $Gini_k$ as reflecting greater need of industry k for the local transport infrastructure is debatable, as previous authors have looked at that variable as capturing a different feature: the degree of product complexity of industry k. In the next two subsections we aim to address more explicitly these two points.

5.1 Endogeneity, Reverse Causation and Waterways Density

Our model has resorted to two critical assumptions that warrant further discussion in case the previous empirical results are intended to be taken as evidence of existence of a *causal* effect from road density to specialisation. Firstly, it has taken r_c as exogenously given. The length of a country's road network is however the result of investment choices in infrastructure, and hence it will respond to a host of economic variables and incentives. Secondly, the model has assumed away any sort of intrinsic differences in productivities *directly* linked to the production functions of final goods laid out in equation (3). In fact, all differences in countries' productivities across final sectors arise *indirectly* from the heterogeneities in the intensity of inputs implied by the parameter α_j.

Relaxing the two above-mentioned assumptions can easily lead to a model where β in (23) can be confounding an effect from road density to specialisation, together with reverse causality from the latter to the former. For example, suppose that for some reason the final good production functions differ across countries, in a way such that H is relatively more productive than F in the final sectors whose α_j lies near one half.[19] In a context like this one, if countries can invest in expanding their road networks, we could well expect r_H to be larger than r_F simply because the incentives to do so are greater in H than in F. From an empirical viewpoint, this reasoning means that β could end up capturing (at least partially) an effect going from patterns of specialisation to road density.

One solution to the above problem would be to find an instrumental variable to generate variation in r_c that is plausibly exogenous. This section does not go that far. However, it intends to provide some further evidence consistent with the main mechanism of the model, relying on a measure of countries' transport network that is *less* sensitive to reverse causality concerns than r_c is. We measure now the internal transport network of an economy by the density of their waterways network. We draw the data on waterways from the CIA World Factbook, and define waterways density as waterways length per square km.[20] Arguably, while countries can still affect the density of their waterways network by investing in creating canals

[19] For example, we may have that final good productions functions are given by (3) for $\alpha_j \in [0, 0.5-\epsilon]$ and for $\alpha_j \in [0.5+\epsilon, 1]$, where $0 < \epsilon < 0.5$, and by

$$Y_j = (1+\phi_c)\frac{1}{\alpha_j^{\alpha_j}(1-\alpha_j)^{1-\alpha_j}}\, X_{0,j}^{1-\alpha_j}\, X_{1,j}^{\alpha_j}, \quad \text{for } \alpha_j \in (0.5-\epsilon, 0.5+\epsilon), \quad \text{with } \phi_H > \phi_F.$$

[20] The CIA World Factbook measures waterways as the total length of navigable rivers, canals and other inland bodies of water.

or improving the navigability of some rivers and bodies of water, the scope for this is far more limited than in the case of building and expanding roads.

One additional aspect we exploit in this section is the possibility that waterways density impacts specialisation *heterogeneously* at different stages of development. For a number of reasons, richer economies tend to have much denser road networks than poorer ones. In particular, poorer economies may find it harder to undertake the necessary investment to build a sufficiently developed road infrastructure. On the other hand, while the presence of waterways may have influenced patterns of development before railroads and roads became more widespread worldwide, waterways are no longer a mode of transportation that seems to greatly impact the current level of development of economies in a systematic way. In fact, a quick look at simple cross-country correlations in Figure 2 shows that, while income per head and road density display a clear positive correlation, the association between income per head and waterways density is rather weak.[21]

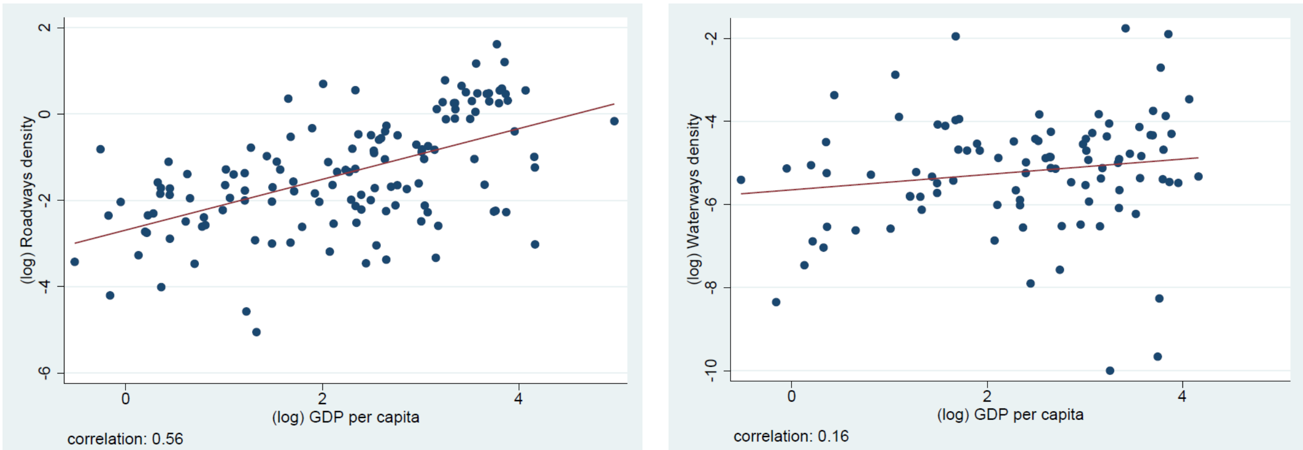


Figure 2: Roadways and waterways density against GDP per head

Table IV displays the results of a regression equation analogous to (23), but where r_c is replaced by a measure of *waterways density*. The table shows the results of two sets of regressions for three different countries sample: entire sample, high-income countries, and low-income countries. The regressions based on the whole set of countries yield an estimate that is negative, and it is also significant. However, this aggregate result masks important heterogeneities

[21] An interpretation of the correlations in Figure 2 is that, as economies grow richer, roadways tend to gradually overshadow waterways as a mode of internal transportation. From this perspective, we could then expect waterways to represent an important determinant of patterns of specialisation in poorer economies, but losing preeminence in richer economies where roadways can more easily make up for an insufficiently deep internal waterway network.

TABLE IV
Waterways Density as Measure of Transport Network

	(1)	(2)	(3)	(4)	(5)	(6)
Waterways Density x Input Narrowness	-0.294**	-0.012	-0.998**	-0.319**	-0.010	-1.292***
	(0.142)	(0.135)	(0.446)	(0.148)	(0.144)	(0.463)
Rule of Law x Input Narrowness	-5.167***	-5.449***	-6.888***	-5.009***	-5.650***	-4.989**
	(1.173)	(1.368)	(2.341)	(1.232)	(1.451)	(2.450)
Fin Dev x Input Narrowness	-0.065***	-0.025*	-0.024	-0.049***	-0.018	0.061
	(0.016)	(0.017)	(0.082)	(0.016)	(0.017)	(0.084)
GDP per capita x Input Narrowness	0.042	0.102	0.053	0.139*	0.208**	0.039
	(0.071)	(0.095)	(0.259)	(0.075)	(0.099)	(0.266)
GDP x Input Narrowness	-0.944***	-0.223	-1.646***	-0.730***	-0.255	-1.692***
	(0.149)	(0.182)	(0.441)	(0.150)	(0.181)	(0.453)
Capital Intensity x $(K/L)_c$				0.157*	-0.095	1.539**
				(0.090)	(0.122)	(0.699)
Skill Intensity x H_c				0.007***	0.020***	-0.020***
				(0.001)	(0.002)	(0.002)
Observations	25,229	13,597	11,632	21,904	11,750	10,154
R-squared	0.773	0.757	0.684	0.810	0.798	0.715
Countries Sample	All	High	Low	All	High	Low
Number of Countries	93	47	46	91	46	45
Number of Industries	294	294	294	259	259	259

Robust standard errors reported in parentheses. The dependent variable is log($Expo_{c,k}$) in year 2014. Waterway density equals internal waterways per square kilometer. Waterways is taken from the CIA World Factbook, and comprises total length of navigable rivers, canals and other inland water bodies. ***p<0.01, **p<0.05, *p<0.1

in the effect of waterways density on export specialisation in the case of richer versus poorer economies. Columns (2) and (5) essentially show that waterways density carries no impact whatsoever in the subsample of above-median income economies. In contrast, columns (3) and (6) exhibit a negative and highly significant coefficient. This result suggests that, in the case of poorer economies, those that enjoy a denser network of waterways tend to export relatively more in industries that require a wider intermediate input base.

5.2 Alternative Interpretations of the Input Breadth Measures

The analysis in Section 4 was based on the notion that the degree of input breadth of industry k can serve as proxy for how reliant this industry is on the internal transport network. The need to source a large variety of inputs can certainly make a particular sector heavily dependent on efficient transportation; however, it can also mean that the sector is highly sensitive to sound contract enforcement. Indeed, Blanchard and Kremer (1997) and Levchenko (2007) have previously used input-output data to compute diversification indices for intermediate input purchases across industries, and use them to proxy the degree of complexity of sectors:

sectors with more diversified (i.e., less concentrated) input bases are considered to be more complex.[22] In their analysis, more complex sectors require better contract enforcement to work efficiently. From this perspective, countries with better functioning institutions should exhibit a comparative advantage in goods that require a wide intermediate input base. Levchenko (2007) shows that this is indeed the case for US imports: the US imports a higher share of goods with greater diversity of intermediate inputs from countries with better rule of law.

Our regressions in Tables II - IV have conditioned on the interaction between rule of law in country c and the Gini coefficient for intermediate inputs in industry k. The estimate of β remained consistently negative and significant, regardless of the introduction of this additional control. In that regard, our results seem to suggest that *both* institutions and local transport networks are instrumental and complementary to the growth of industries with wide input bases. This section will attempt to further strengthen this argument

Countries with better institutions are in general richer, and also exhibit a denser transportation infrastructure network. If $Gini_k \times r_c$ in (23) were *not* capturing any type of impact related to how transport-intensive industry k is, but *only* the effect of rule of law in country c through its correlation with r_c, then the correlation found in Table I should arise more prominently for industries that are relatively more dependent on judicial quality. The regressions reported in Table V show this is actually not the case in the data.

Columns (1) and (2) in Table V show the results of the simple correlation reported initially in column (1) of Table I, after splitting the set of industries in two subsamples: low contract intensity vs. high contract intensity. To do so, we take the measure of contract intensity by industries from Nunn (2007), and split the sample of industries according to whether they rank below or above the median value of contract intensity.[23] If the $Gini_k$ were simply proxying for how sensitive to efficient contract enforcement industry k is, then the estimate in column (1) should turn out to be significantly milder than that one in column (2). The regressions show, however, that the negative correlation is significant in both subsamples, and moreover they are of very similar magnitude.

Column (3) carries out a similar assessment, but instead of splitting the sample of industries in two subsets, it introduces a triple interaction term between r_c, $Gini_k$, and the measure of contract intensity of industry k. As we can observe from the table, the coefficient associated to the triple interaction is not statistically significant.[24]

[22] Both articles used the Herfindahl index of concentration instead of the Gini as their benchmark measure.

[23] Nunn (2007) reports contract intensity measures for 222 industries coded according to NAICS 1997. We lose some of the original industries in Table I when matching the NAICS 1997 codes to those of BAE.

[24] In column (3), if $Gini_k$ in were not capturing any type of impact related to how transport-intensive industry

TABLE V

Regressions on Industry Subsamples: Effects ar Different Leves of Contract Intensity

	(1)	(2)	(3)	(4)	(5)
Road Density x Input Narrowness	-5.034***	-4.147***	-3.831**	-3.776***	-2.753***
	(0.991)	(0.813)	(1.643)	(0.700)	(0.686)
Rule of Law x Contract Intensity				0.262***	0.167***
				(0.061)	(0.061)
Road Dens x Input Nwness x Contract Intensity			-1.007		
			(2.781)		
Road Density x Contract Intensity			1.275		
			(2.627)		
Rule of Law x Input Narrowness				1.212	-0.139
				(1.102)	(1.148)
Fin Dev (priv cred) x Input Narrowness				-0.040**	-0.014
				(0.018)	(0.019)
GDP per capita x Input Narrowness				-0.046	-0.053
				(0.052)	(0.050)
GDP x Input Narrowness				-0.881***	-0.718***
				(0.165)	(0.170)
Skill Intensity x H_c					0.007***
					(0.001)
Capital Intensity x $(K/L)_c$					0.346***
					(0.087)
Observations	13,192	13,056	26,248	25,476	20,984
R-squared	0.718	0.825	0.762	0.765	0.802
Contract Intensity	low	high	all	all	all
Number of Industries	97	96	193	193	172
Number of Countries	135	136	136	136	122

Robust standard errors reported in parentheses. All regressions include country and industry fixed effects. The dependent variable is log($Expo_{c,k}$) in year 2014. Contract intensity by industry is taken from Nunn (2007). The original measures are coded on the NAICS 1997 classification, and matched to the BAE codes. Nunn (2007) measures contract intensity in k as the proportion of inputs of k classified as differentiated by Rauch (1999). *** p<0.01, ** p<0.05, * p<0.1

Finally, columns (4) and (5) report the results of two regressions analogous, respectively, to those in columns (3) and (4) of Table II, but including also the interaction term between rule of law in country c and contract intensity of industry k. Consistently with the previous results in the literature, the regressions show that countries with better institutions exhibit a comparative advantage in the industries with high levels of contract intensity. In addition to that, the regressions still support the model's prediction that countries with denser road networks export relatively more in those sectors that need to source a larger variety of intermediate inputs.

k is, but only an indirect correlation between r_c and institutions, then the coefficient associated to the triple interaction term should have a positive and significant estimate.

6 Concluding Remarks

We have proposed a simple trade model where the density of the internal transport network represents a key factor in shaping comparative advantage and specialisation. The underlying mechanism rests on the idea that moving intermediate inputs across spatial locations is costly. As a consequence, industries whose production processes rely on a wide set of intermediate inputs become heavier users of the internal transport network. The model shows that countries with denser transport infrastructures exhibit a comparative advantage on industries that combine a large variety intermediate goods. Furthermore, when disparities in the internal transport network across countries are sufficiently pronounced, they can sometimes even overturn more standard Ricardian patterns of specialisation driven by heterogeneities in labour productivities across sectors.

Drawing on intermediate goods transactions from the US input-output matrix to measure industries' input breadth, we have also shown that the patterns of specialisation predicted by the model are broadly consistent with the trade flows observed in the data. In particular, our empirical analysis shows that countries with denser road networks tend to export relatively more in industries that rely on a wide set of intermediate inputs.

Several caveats apply nevertheless to the empirical evidence. First, patterns of specialisation could be influencing investment in road infrastructure, and thus be behind the correlation found in the data. In that respect, the fact that the same correlation appears when substituting roadway density by waterway density seems quite reassuring, as waterways are much more difficult to expand in response to increased transport needs than road networks. Second, our measure of input breadth by industries could alternatively be capturing a stronger need for contract enforcement when the input base is wider. We showed however that the correlation predicted by our model is still present when the confounding effect of institutional quality (by country) and judiciary intensity (by industry) is also taken into account. Finally, our measure of road density is a relatively imprecise way to capture the efficiency in connectedness of different locations within a country. Road networks not only differ in length, but they also differ in terms of width, surface quality, driving speed they can safely allow, etc. Furthermore, our measure of road density also fails to account for inefficiencies in the layout of internal roads. It would be certainly desirable to use of a more detailed and encompassing measure of efficiency of internal road networks by countries. Yet, unless the above sources of measurement error are somehow more prevalent in economies specialised in industries with wide input bases, it would be hard to argue that they can end up largely distorting the empirical results in a systematic way.

Appendix A: Proofs and Additional Theoretical Results

Proof of Lemma 1. To prove that $c_j^*(r_1)/c_j^*(r_2) \geq 1$ for all $\alpha_j \in [0,1]$ with strict inequality if and only if $\alpha_j \in (0,1)$, notice that $\varphi'(r) < 0$ implies that $1+\varphi(r_1)t > 1+\varphi(r_2)t$, and using (7) we have:

$$c_j^*(r_1)/c_j^*(r_2) = \begin{cases} \left[(1+\varphi(r_1)t) \,/\, (1+\varphi(r_2)t)\right]^{\alpha_j} & \text{for } 0 \leq \alpha_j \leq 0.5, \\ \left[(1+\varphi(r_1)t) \,/\, (1+\varphi(r_2)t)\right]^{1-\alpha_j} & \text{for } 0.5 \leq \alpha_j \leq 1. \end{cases} \tag{24}$$

To prove the second part of the lemma, apply logs to $c_j^*(r_1)/c_j^*(r_2)$ in (24), and differentiate w.r.t. α_j, to obtain that $\partial \left(\ln \left(c_j^*(r_1)/c_j^*(r_2)\right)\right)/\partial\alpha_j > 0$ for $0 \leq \alpha_j < 0.5$ and $\partial \left(\ln \left(c_j^*(r_1)/c_j^*(r_2)\right)\right)/\partial\alpha_j < 0$ for $0.5 < \alpha_j \leq 1$. ■

Proof of Proposition 1. We first prove by contradiction that $\omega = 1$ cannot hold in equilibrium. Given that the expression in (16) entails that total imports from F by H increase with ω, while total exports by H to F decrease with ω, it will then follow that in equilibrium we must necessarily have $\omega^* > 1$, and that this equilibrium will be unique. We carry out the proof of $\omega^* > 1$ by splitting the possible parametric configurations of the model in three subsets.

i) Case 1: $\delta \leq (1+\tau)^2$. In this case, when $\omega = 1$, using the LHS of (16), it follows that total exports by H are equal to:

$$Expo_H = \frac{\ln(1+\varepsilon) - \ln(1+\tau)}{2\ln(1+\varepsilon) - \ln(\delta)}. \tag{25}$$

Notice that $(1+\tau)^2 \geq \delta$ implies the RHS of (25) is never greater than one half, while Assumption 2 implies it is strictly above zero. Using now the RHS of (16), we can obtain that total imports by H are:

$$Impo_H = 1 - \frac{\ln(1+\varepsilon) + \ln(1+\tau) + \ln(\delta)}{2\ln(1+\varepsilon) + \ln(\delta)}. \tag{26}$$

Comparing (25) versus (26), while bearing in mind $\delta > 1$, yields $Expo_H > Impo_H$. Hence, when $(1+\tau)^2 \geq \delta$, the equilibrium must necessarily encompass $\omega > 1$.

ii) Case 2: $(1+\tau)^2 < \delta < (1+\varepsilon)^2$. Using setting again (16), we obtain:

$$Expo_H = \frac{\ln(1+\varepsilon) - \ln(1+\tau) + \ln(\delta)}{2\ln(1+\varepsilon) + \ln(\delta)}, \tag{27}$$

while total imports by H are still given by (26). When $(1+\tau)^2 < \delta$, the RHS of (27) yields a value strictly larger than one half, while the RHS of (26) is always strictly smaller than one half. As a consequence, $Expo_H > Impo_H$ also when $(1+\tau)^2 < \delta < (1+\varepsilon)^2$, and the equilibrium must necessarily encompass $\omega > 1$ in that range too.

iii) Case 3: $(1+\varepsilon)^2 < \delta$. Using once again (16), notice that total exports by H are still given by (27), which yields a value strictly above 0.5 and strictly below 1. In addition, total imports by H are still given by (26), which yields a value strictly above 0, but strictly below one half. Hence, when $(1+\tau)^2 \geq \delta$, the equilibrium must encompass $\omega > 1$ as well.

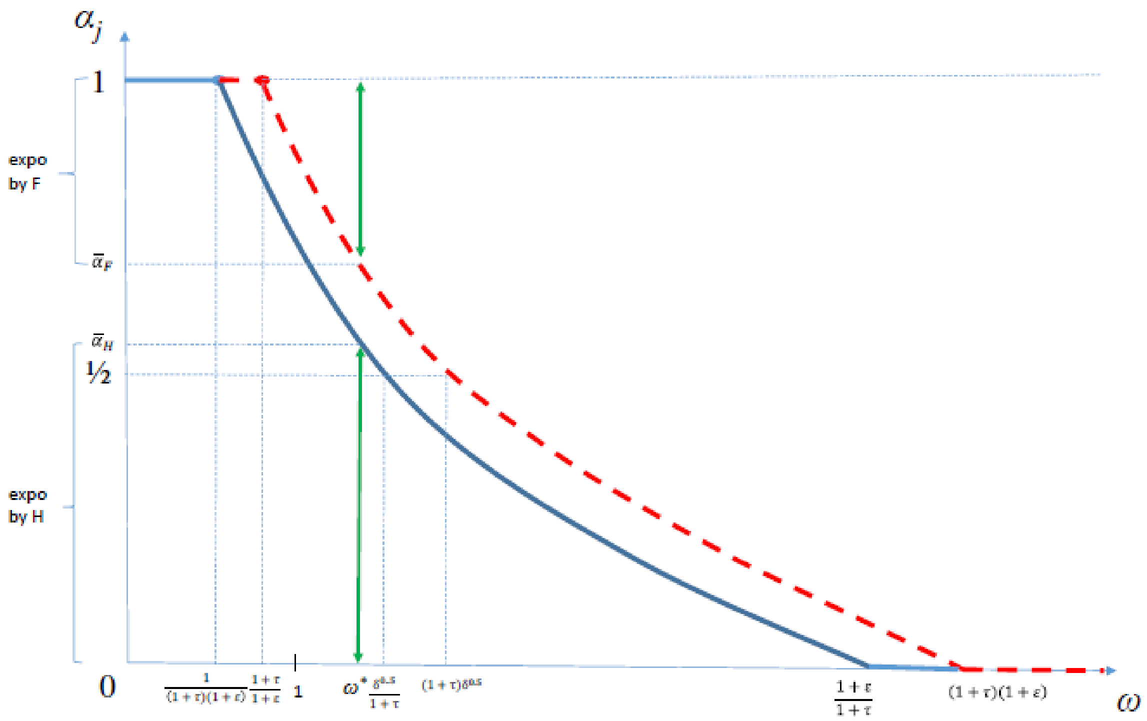
Next, to prove that ω^* is strictly increasing in δ, it suffices to note that the boundaries in all the expressions of the indicator functions $\mathbb{I}(\cdot)$ in (16) are all strictly increasing in δ.

Lastly, to prove the different bounds on ω^* we proceed by contradiction for each of them. First, suppose that $(1+\varepsilon)^2 > \delta$. Notice that if $\omega^* \geq (1+\tau)\delta^{0.5}$, then using (16) we can observe that the mass of final goods exported by F would be at least one half. However, this is incompatible with the fact that in equilibrium $\omega^* > 1$. Hence, it must be that $\omega^* < (1+\tau)\delta^{0.5}$. Next, notice that when $\omega^* \geq (1+\varepsilon)/(1+\tau)$, the exports by H fall to zero, while H's imports are strictly positive; hence, this cannot hold in equilibrium either, and it must be that $\omega^* < (1+\varepsilon)/(1+\tau)$. Second, suppose now that $(1+\varepsilon)^2 < \delta$. In this case when $\omega^* \geq (1+\tau)^{-1}\delta^{0.5}$, exports by H would fall to zero, while H's imports would still be strictly positive. Hence, it must be the case that $\omega^* < (1+\tau)^{-1}\delta^{0.5}$. ■

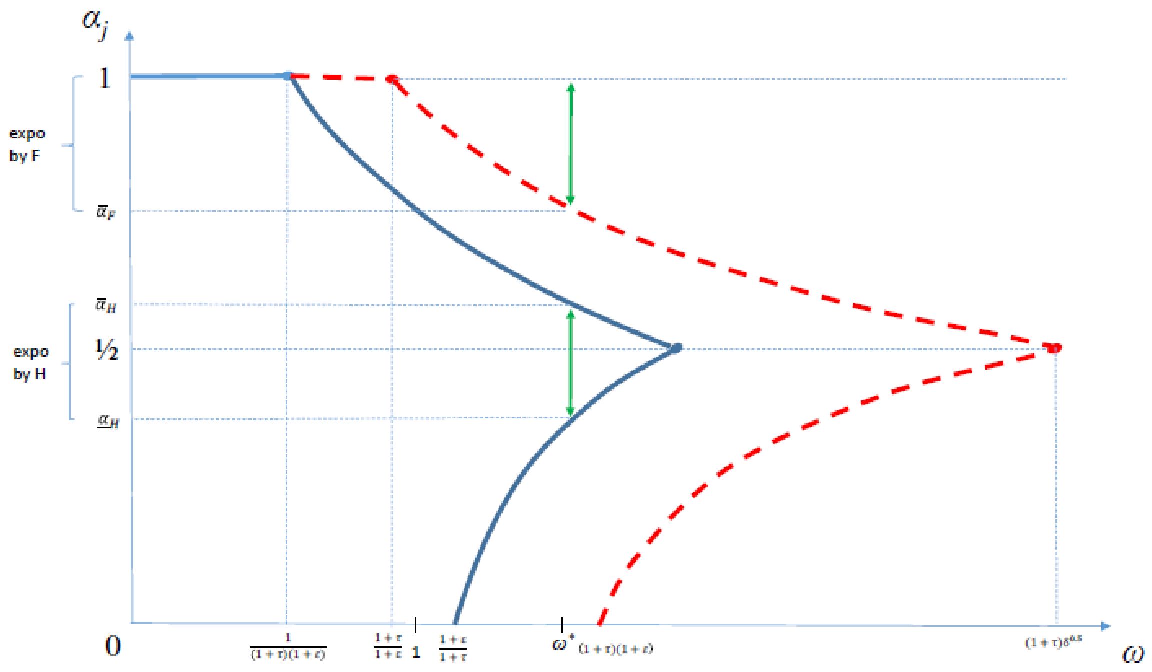
Additional Results of Proposition 2. We plot below that additional two sub-cases encompassed by Proposition 2 – *case 1 (bis)* and *case 2 (bis).*

Case 1 (bis) occurs when $\delta < (1+\varepsilon)^2$ and the parametric configuration of the model is such that, in equilibrium, $\omega^* < (1+\tau)^{-1}\delta^{\frac{1}{2}}$. The graph is qualitatively similar to case 1 in Figure 1 of the main text, with the difference that country H exports goods with $\alpha_j \in [0, \overline{\alpha}_H)$, where $\overline{\alpha}_H$ lies above one half.

Case 2 (bis) takes place when $\delta > (1+\varepsilon)^2$ and the parametric configuration of the model is such that, in equilibrium, $\omega^* > (1+\varepsilon)(1+\tau)$. The main qualitative difference between this one and case 2 in Figure 1 is that here country F exports only those goods on the upper end of $[0,1]$, namely those with $\alpha_j \in (\overline{\alpha}_F, 1]$.



case 1 (bis): $\boldsymbol{\delta < (1+\varepsilon)^2}$ **and** $\boldsymbol{\omega^* < (1+\tau)^{-1}\gamma^{-½}}$



case 2 (bis): $\boldsymbol{\delta > (1+\varepsilon)^2}$ **and** $\boldsymbol{\omega^* < (1+\tau)(1+\varepsilon)}$

Figure 1 (bis): Patterns of Trade and Specialisation

A Simple Microfoundation of the Distance Function $\varphi(r)$

We provide here a simple illustration of the geographical structure an the economy that serves as microfoundation of the function $\varphi(r)$ assumed in Subsection 2.2.

Suppose there exists an infrastructure network connecting location 0 and 1 that comprise two two types of pathways. One is a semi-circular path of total length $\pi/2$, which represents the *least* direct path between the two locations. The other one is a road of length $r \in [0, 1]$, which allows shortening the distance between the two locations given by the semi-circular path.

The infrastructure network structure is plotted in Figure 3 for two different roads lengths, namely $0 < r_1 < r_2 < 1$. (Without any loss of generality, we will arbitrarily place the starting point of the roads always in location 0.) The two extreme cases $r = 0$ and $r = 1$ would correspond, respectively, to the case where the only path available from location 0 to location 1 is via the semi-circular arch of length $\pi/2$, and the case where the two locations are connected by a straight horizontal line of length one.[25]

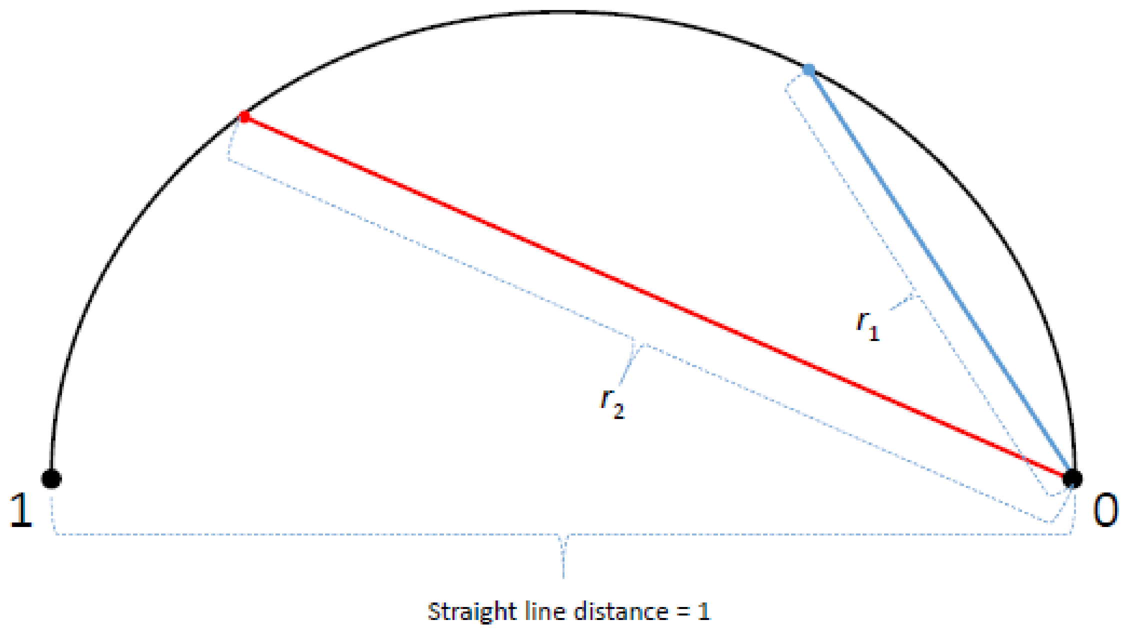


Figure 3: Geographic Structure of the Economy

[25] Conceptually, Figure 3 intends to represent the notion that road networks facilitate transportation across location 0 and 1, relative to the (lengthier) semi-circular path of length $\pi/2$. Nothing precludes the fact that a road network of length r could comprise several segments, whose lengths sum up to r. As it will become clear next, this would not be optimal. More precisely, given a total length of road network equal to $r \in [0, 1]$, the shortest distance to connect inputs locations is achieved by building one single straight line.

Denoting by $\varphi(r)$ the *shortest* distance between location 0 and 1, given a road network of length $r \in [0,1]$. The following lemma characterises the main properties of $\varphi(r)$.

Lemma 3 *The shortest distance between location* 0 *and* 1 *is a strictly decreasing function of the road length; i.e.,* $\varphi'(r) < 0$*. In addition,* $\varphi''(r) < 0$.[26]

Proof. To prove the lemma we first find the closed-form expression of the function $\varphi(r)$. Notice that any straight segment connecting two points of the semi-circle linking location 0 to location 1 could be seen as a chord within a circle. We can then use the formula for the *chord length* in circle, which in this case of only one straight segment of length r would state that

$$r = 2 \cdot radius \cdot \sin\left(f \cdot \frac{\pi}{2}\right), \tag{28}$$

where f equals the share of the whole semi-circle that is covered by the straight line of length r.[27] From (28), bearing in mind that $radius = 0.5$, it follows that the share of the semi-circle covered by the road of length r is given by $f = \frac{2}{\pi} \cdot \arcsin(r)$. Therefore, since the semi-circle has total length equal to $\pi/2$, the total distance *not* covered by r is equal to $\pi/2 - \arcsin(r)$. Summing up to this last amount the total length of the road, r, we finally obtain that:

$$\varphi(r) = \frac{\pi}{2} + r - \arcsin(r), \tag{29}$$

Next, differentiating (29), we obtain

$$\varphi'(r) = 1 - (1 - r^2)^{-\frac{1}{2}}, \tag{30}$$

which is strictly negative for any $0 < r \leq 1$.

Finally, differentiating (30) , we can also observe that $\varphi''(r) = -r/(1 - r^2)^{-\frac{3}{2}}$, which is also strictly negative for any $0 < r \leq 1$. Finally, notice that $\varphi''(r) < 0$ in turn implies that the shortest way to link location 0 and 1 when the road length is $r \in [0,1]$ is through one single straight segment of length r. ■

---

[26] The *only* crucial feature that the model needs is $\varphi'(r) < 0$, which implies that road length lowers the cost of internal transportation of goods. This is in fact the only assumption placed in the main text in Section 2.2. The only implication of $\varphi''(r) < 0$ for the model is that, for a given road length r, the shortest distance between location 0 and 1 is achieved by building one single straight segment of length r, as plotted in Figure 3.

[27] More formally, f equals the ratio between the angle formed by a straight line going from the centre of the semi-circle to the endpoint of the chord of length r, and an angle of 180°.

Appendix B: Additional Empirical Results

This appendix displays some additional results, which are complementary to those in Section 4.3, including some summary statistics presented in Table A.1. Next, in Table A.2 and A.3 show the results of some regressions, analogous to some of those previously presented in Section 4.3, but where the measure of transport network in country is either changed or expanded.

TABLE A.1

Summary Statistics

Variable	Mean	Std Dev	Min	Max	Obs
Gini Coef.	0.93949	0.02314	0.88882	0.99342	294
Herfindahl Index	0.10087	0.08036	0.02883	0.77590	294
Coef. Var.	5.8063	2.03478	3.1955	17.2480	294
Log-Variance	-8.4915	0.62126	-9.5830	-6.2111	294
Roadways per sq km	0.54401	0.73987	0.00639	5.04494	137
Railways per sq km	0.02241	0.02883	0.00007	0.13693	128
Waterways per sq km	0.01334	0.02740	0.00005	0.17264	98

In Table A.2, road density is replaced by railway density, as our main measure of depth of local transport network. As we can see, all the results follow a similar pattern as those previously obtained in Section 4.3.

TABLE A.2

Transport Density measured by Railway Density

	(1)	(2)	(3)	(4)
Railway Density x Input Narrowness	-2.263***	-1.366***	-1.039***	-1.014***
	(0.133)	(0.161)	(0.163)	(0.164)
Rule of Law x Input Narrowness		-2.686***	-3.072***	-3.073***
		(0.999)	(1.086)	(1.086)
Fin Dev x Input Narrowness		-0.078***	-0.050***	-0.049***
		(0.015)	(0.016)	(0.016)
GDP per capita x Input Narrowness		0.096	0.142**	0.173***
		(0.061)	(0.064)	(0.064)
GDP x Input Narrowness		-0.980***	-0.787***	-0.797***
		(0.145)	(0.149)	(0.151)
Capital Intensity x $(K/L)_c$			0.192**	
			(0.082)	
Skill Intensity x H_c			0.007***	
			(0.001)	
Observations	32,414	31,468	26,204	26,204
R-squared	0.752	0.753	0.794	0.794
Number of Countries	119	115	108	108
Number of Industries	294	294	259	259

Robust standard errors in parentheses. All regressions include country and industry fixed effects. Railway density equals the total length of the railway network in km, divided by the area measured in sq km. Data of railway network length is taken from the CIA factbook. *** p<0.01, ** p<0.05, * p<0.1

In Table A.3, we expand the measure of transport network density to include, in addition to roadways, also railways and waterways. In this case, the density of the transport network of country c is measured as the sum of total kilometers of roadways, railways and waterways, divided by the area of the country. For many countries in the dataset, we do not have information on railways or waterways, while we do have information on roadways. This means that the total sample of the regressions in columns (1), (2) and (3) falls substantially relative to their relevant counterparts in Section 4.3. For additional comparison, in columns (4), (5) and (6), we also include countries with missing information on either railways or waterways (or in both), replacing the missing values by zeros. Again, all the results follow a similar qualitative pattern as those in Section 4.3.

TABLE A.3

Transport density measured by sum of roadways, railways and waterways per square km

	(1)	(2)	(3)	(4)	(5)	(6)
Transp. (road + rail + waterway) Density x Input Nwness	-7.241***	-3.117***	-1.844***	-7.295***	-3.841***	-2.566***
	(0.520)	(0.602)	(0.588)	(0.484)	(0.554)	(0.550)
Rule of Law x Input Narrowness		-3.839***	-4.415***		-1.516*	-1.829**
		(1.227)	(1.293)		(0.887)	(0.947)
Fin Dev (priv cred) x Input Narrowness		-0.073***	-0.057***		-0.083***	-0.047***
		(0.016)	(0.016)		(0.015)	(0.016)
GDP per capita x Input Narrowness		0.077	0.161**		0.004	0.022
		(0.072)	(0.076)		(0.039)	(0.039)
GDP x Input Narrowness		-0.934***	-0.718***		-0.909***	-0.705***
		(0.145)	(0.149)		(0.144)	(0.148)
Capital Intensity x $(K/L)_c$			0.173**			0.221***
			(0.091)			(0.080)
Skill Intensity x H_c			0.007***			0.007***
			(0.001)			(0.001)
Observations	24,912	23,966	20,774	36,069	35,123	29,062
R-squared	0.768	0.769	0.806	0.755	0.756	0.797
Number of Countries	91	87	85	136	132	122
Number of Industries	294	294	259	294	294	259

Robust standard errors in parentheses. All regressions include country and industry fixed effects. Columns (1) to (3) only include observations where information on all transport measures (i.e., roadway, railway and waterway) is available. Columns (4) to (6) also include observations where information on either railway or waterway (or both) are missing, replacing the missing values by zero. *** p<0.01, ** p<0.05, * p<0.1

References

[1] Acemoglu, D., Antràs, P. and Helpman, E. (2007). 'Contracts and Technology Adoption', *American Economic Review*, vol. 97, pp. 916-943.

[2] Acemoglu, D. and Zilibotti, F. (1997). 'Was Prometheus Unbound by Chance? Risk, Diversification and Growth', *Journal of Political Economy*, vol. 105, pp. 709-751.

[3] Acemoglu, D. and Zilibotti, F. (1999). ' Information Accumulation in Development', *Journal of Economic Growth*, vol. 4, pp. 5-38.

[4] Agnosteva, D., Anderson, J. and Yotov, Y. (2014). 'Intra-national Trade Costs: Assaying Regional Frictions', NBER WP No. 19872.

[5] Allen, T. and Arkolakis, C. (2014). 'Trade and the Topography of the Spatial Economy', *Quarterly Journal of Economics*, vol. 129, pp. 1085–1140.

[6] Anderson, J. E., and Van Wincoop, E. (2004). 'Trade costs', *Journal of Economic Literature*, vol. 42, pp. 691–751.

[7] Antràs, P. (2005). 'Incomplete Contracts and the Product Cycle', *American Economic Review,* vol. 95, pp. 1054-1073.

[8] Atkin, D. and Donaldson, D. (2015). 'Who's Getting Globalized? The Size and Implications of Intra-national Trade Costs', Mimeo, MIT.

[9] Barro, R. and Lee, J. (2013). 'A New Data Set of Educational Attainment in the World, 1950-2010', *Journal of Development Economics,* vol 104, pp.184-198.

[10] Beck, T. (2002). 'Financial Development and InternationalTrade. Is There a Link?', *Journal of International Economics*, vol. 57, 107-131.

[11] Becker, B., Chen, J. and Greenberg, D. (2012). 'Financial development, fixed costs, and international trade', *Review of Corporate Finance Studies*, vol. 2, pp. 1-28.

[12] Becker, R., Gray, W. and Marvakov, J. (2013). 'NBER-CES manufacturing industry database: Technical notes', NBER.

[13] Blanchard, O. and Kremer, M. (1997). 'Disorganization', *Quarterly Journal of Economics*, Vol. 112, pp. 1091-1126.

[14] Coşar, K. and Fajgelbaum, P. (2016). 'Internal Geography, International Trade, and Regional Specialization', *American Economic Journal: Microeconomics*, vol. 8, pp. 24-56.

[15] Costinot, A. (2009), 'On the Origins of Comparative Advantage', *Journal of International Economics*, vol. 77, pp 255-264.

[16] Donaldson, D. (forthcoming). 'Railroads of the Raj: Estimating the Economic Impact of Transportation Infrastructure', *American Economic Review.*

[17] Donaldson, D. and Hornbeck, R. (2016). 'Railroads and American Economic Growth: A "Market Access" Approach', *Quarterly Journal of Economics*, vol. 131, pp 799-858.

[18] Dornbusch, R., Fischer, S. and Samuelson, P. (1977). 'Comparative Advantage, Trade, and Payments in a Ricardian Model with a Continuum of Goods', *American Economic Review*, vol. 67, pp. 823-839.

[19] Duranton, G., Morrow, P. and Turner, M. (2014). 'Roads and Trade: Evidence from the US', *Review of Economic Studies*, vol. 81, pp. 681-724.

[20] Fajgelbaum, P. and Redding, S. (2014). 'External Integration, Structural Transformation and Economic Development: Evidence from Argentina 1870-1914', NBER WP No. 20217.

[21] Gaulier, G. and Zignago, S. (2010). 'BACI: International Trade Database at the Product-level. The 1994-2007 Version', CEPII Working Paper 2010-23.

[22] Greenwood, J. and Jovanovic, B. (1990). 'Financial Development, Growth, and the Distribution of Income', *Journal of Political Economy*, vol. 98, pp. 1076-1107.

[23] Hillberry, R. and Hummels, D. (2008). 'Trade responses to geographic frictions: A decomposition using micro-data', *European Economic Review*, vol. 52, pp. 527-550.

[24] Imbs, J. and Wacziarg, R. (2003). 'Stages of Diversification', *American Economic Review*, vol. 93, pp. 63-86.

[25] Levchenko, A. (2007). 'Institutional Quality and International Trade', *Review of Economic Studies*, vol. 74, pp. 681-724.

[26] Limao, N. and Venables, A. (2001). 'Infrastructure, geographical disadvantage, transport costs, and trade', *World Bank Economic Review*, vol. 15, pp. 451–479.

[27] Manova, K. (2013). 'Credit Constraints, Heterogeneous Firms, and International Trade', *Review of Economic Studies*, vol. 80, pp. 711-744.

[28] Matsuyama, K. (2017). 'Geographical Advantage: Home Market Effect in a Multi-Region World', *Research in Economics*, vol. 71, pp. 740-758.

[29] Mesquita Moreira, M., Blyde, J., Volpe, C. and Molina, D. (2013). 'Too Far to Export: Domestic Transport Costs and Regional Export Disparities in Latin America and the Caribbean', Washington, DC, IADB.

[30] Nunn, N. (2007). 'Relationship-specificity, incomplete contracts, and the pattern of trade', *Quarterly Journal of Economics*, vol. 122, pp.569-600.

[31] Ramondo, N., Rodriguez-Clare, A. and Saborio-Rodriguez, M. (2016). 'Trade, Domestic Frictions, and Scale Effects', *American Economic Review*, vol. 106, pp. 3159-84.

[32] Rauch, J. (1999). 'Networks versus markets in international trade', *Journal of International Economics*, vol. 48, pp. 7-35.

[33] Redding, S. (2016). 'Goods Trade, Factor Mobility and Welfare', *Journal of International Economics*, vol. 101, pp. 148-167.

[34] Saint-Paul, G. (1992), 'Technological Choice, Financial Markets and Economic Development', *European Economic Review*, vol. 36, pp. 763-781.

[35] Svaleryd, H. and Vlachos, J. (2005) 'Financial markets, the pattern of industrial specialization and comparative advantage: Evidence from OECD countries', *European Economic Review*, vol. 49, pp. 113-144.

[36] Volpe Martincus, C. and Blyde, J. (2013). 'Shaky roads and trembling exports: Assessing the trade effects of domestic infrastructure using a natural experiment', *Journal of International Economics*, vol. 90, pp. 148-161.

[37] Volpe Martincus, C., Carballo, J. and Cusolito, A. (2017). 'Roads, exports and employment: Evidence from a developing country', *Journal of Development Economics*, vol. 125, pp. 21-39.